GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2002, 19:02:47 ; Search time 2325.7 Seconds
(without alignments)
992.189 Million cell updates/sec

Title: US-09-303-518D-125
Perfect score: 1344
Sequence: 1 atgattaaaatcaaaaaagg.....ccattgagaaggaaggctga 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1344	100.0	1344	20	AAZ12026	Neisseria meningit
c	2	1344	100.0	44608	21	AAA81495	N. meningitidis pa
c	3	1344	100.0	349980	21	AAF21607	Neisseria meningit
	4	1229.4	91.5	1344	20	AAZ12027	Neisseria meningit
	5	1212.8	90.2	1344	20	AAZ12028	Neisseria gonorrho
	6	521.8	38.8	1830121	17	AAT42063	Haemophilus influe
	7	473.6	35.2	474	20	AAZ12025	Neisseria meningit
	8	473.6	35.2	474	21	AAA81335	N. meningitidis Me
c	9	363	27.0	363	21	AAZ54034	Neisseria meningit
c	10	327.8	24.4	363	21	AAZ54035	Neisseria meningit
c	11	326.4	24.3	96109	22	AAF28548	Genomic fragment #
c	12	324.6	24.2	363	21	AAZ54033	Neisseria gonorrhe
	13	113.6	8.5	1353	20	AAX91657	Porphorymonas ging
	14	113.6	8.5	1362	20	AAX91536	Porphorymonas ging
c	15	40.8	3.0	5059	20	AAX84332	Stealth virus nucl
	16	38.2	2.8	38734	20	AAZ32020	Human METH1 relate
	17	38.2	2.8	38734	22	AAC90077	AL021529 cDNA clon
	18	37.8	2.8	1416	23	AAS51578	Pseudomonas aerugi
	19	37.8	2.8	1539	21	AAC45663	Arabidopsis thalia
	20	37.8	2.8	1541	21	AAC33763	Arabidopsis thalia
c	21	37.6	2.8	92934	21	AAA81473	N. meningitidis pa
c	22	37.6	2.8	172325	21	AAF21613	Neisseria meningit
c	23	37.6	2.8	837096	21	AAA81489	N. meningitidis pa
	24	37	2.8	1980	21	AAZ51001	X. oryzae avrRxo r
	25	37	2.8	3720	21	AAZ51000	X. oryzae avrRxol
	26	36.4	2.7	1401	21	AAZ95378	Chlamydia pneumoni
c	27	36.4	2.7	1725	22	AAC85030	Human Zcytor14 var
c	28	36.4	2.7	2076	22	AAC85028	Human cytokine rec
c	29	36.4	2.7	2109	24	AAS18131	Human DCRS7 revers
	30	35.6	2.6	125401	22	AAD17186	Streptomyces nours
	31	35.4	2.6	296	23	AAS57519	cDNA #195 encoding
c	32	35.4	2.6	763	20	AAZ09101	Mistletoe lectin A
c	33	35.4	2.6	1598	20	AAZ09100	Mistletoe lectin D
	34	35.4	2.6	9810	20	AAZ32025	Human METH1 relate
	35	35.4	2.6	9810	22	AAC90082	AF018073 cDNA clon
	36	35	2.6	323	23	AAS86095	DNA encoding novel
	37	35	2.6	1139	21	AAF21735	Human breast and o
	38	35	2.6	1689	23	AAS86098	DNA encoding novel
	39	35	2.6	4383	7	AAN60550	Mouse erythroid Ba
	40	34.8	2.6	1325	22	AAD08837	Human G-protein co
c	41	34.6	2.6	697	22	AAF25080	Nucleotide sequenc
c	42	34.6	2.6	828	22	AAF88318	S. spinosa DNA fra
	43	34.6	2.6	45624	22	AAF88315	S. spinosa DNA fra
c	44	34.6	2.6	50000	22	AAF88312	S. spinosa DNA fra
c	45	34.6	2.6	80161	20	AAZ21501	DNA fragment of Sa
c	46	34.4	2.6	762	20	AAZ09104	Mistletoe lectin A
	47	34.4	2.6	1677	22	AAF61095	P. putida KT2440-a
	48	34.2	2.5	927	21	AAZ53601	Neisseria gonorrhe
c	49	34.2	2.5	1680	23	AAS54169	Pseudomonas aerugi
c	50	34.2	2.5	1968	21	AAC96946	Human secreted pro
	51	34.2	2.5	4048	23	ABL02092	Drosophila melanog
	52	34.2	2.5	4242	23	ABL03672	Drosophila melanog
	53	34.2	2.5	7732	20	AAX77723	pDV2 vector. Synt
c	54	34	2.5	5292	21	AAA35252	Human adenosine re
c	55	33.8	2.5	25365	23	AAS59558	Propionibacterium
	56	33.6	2.5	2206	15	AAQ70754	Beta tubulin gene
	57	33.6	2.5	3445	14	AAQ48230	Acremonium chrysog
	58	33.6	2.5	3445	14	AAQ55405	A.chrysogenum beta
	59	33.6	2.5	3445	14	AAQ55406	A.chrysogenum beta
	60	33.6	2.5	4451	19	AAV23288	Synthetic human Fa
	61	33.6	2.5	6761	21	AAA57362	DNA encoding a hum
	62	33.6	2.5	4411529	22	AAI99682	Mycobacterium tube
	63	33.4	2.5	10198	22	AAF61282	N. magadaii bacter
	64	33.2	2.5	963	22	AAS41208	cDNA encoding nove
	65	33.2	2.5	1020	23	AAS76925	DNA encoding novel
c	66	33.2	2.5	1050	23	AAS73416	DNA encoding novel
c	67	33.2	2.5	1050	23	AAS80144	DNA encoding novel
c	68	33	2.5	545	21	AAF15192	Trichoderma reesei
	69	33	2.5	609	21	AAZ54464	Neisseria gonorrhe
	70	33	2.5	609	21	AAZ54467	Neisseria gonorrhe
	71	33	2.5	783	23	AAS54375	Pseudomonas aerugi
c	72	33	2.5	1380	20	AAZ12065	Neisseria gonorrho
	73	33	2.5	1596	17	AAT03695	Beta-amylase S291A
	74	33	2.5	6312	17	AAT03696	Plasmid pBETA92, e
	75	33	2.5	10732	21	AAA10594	Gene encoding a su
	76	33	2.5	44377	18	AAT78508	Platenolide syntha
	77	33	2.5	44377	18	AAT80414	Platenolide syntha
	78	32.8	2.4	408	22	AAH52101	Human AFP protein
c	79	32.8	2.4	2126	22	AAH14849	Human cDNA sequenc
c	80	32.8	2.4	2348	20	AAZ07116	Human vanilloid re
c	81	32.8	2.4	2351	20	AAZ07114	Human vanilloid re
c	82	32.8	2.4	2380	20	AAX87492	Human vanilloid re

c	83	32.8	2.4	2469	22	AAC60297	Human vanilloid re
c	84	32.8	2.4	2765	21	AAA14874	DNA encoding a van
c	85	32.8	2.4	2779	19	AAV59691	Human secreted pro
c	86	32.8	2.4	2783	20	AAZ22829	Human vanilloid re
c	87	32.8	2.4	2809	21	AAA30254	Human VR-2 coding
c	88	32.8	2.4	2825	22	AAD09554	Human transporter
c	89	32.8	2.4	3224	23	AAS82583	DNA encoding novel
c	90	32.8	2.4	3224	23	AAS89304	DNA encoding novel
	91	32.8	2.4	4134	23	AAS89154	DNA encoding novel
	92	32.8	2.4	27541	22	AAD17185	Streptomyces nours
c	93	32.6	2.4	547	22	AAS05581	Mammalian vestibul
	94	32.6	2.4	927	21	AAZ53603	Neisseria meningit
	95	32.6	2.4	1116	22	AAF28683	Human protein HP03
	96	32.6	2.4	1116	24	ABA02176	Human G protein-co
	97	32.6	2.4	1119	21	AAA64367	DNA encoding a P2Y
	98	32.6	2.4	1119	21	AAD01119	Human orphan G pro
	99	32.6	2.4	1119	21	AAA46018	Human G protein co
	100	32.6	2.4	1119	22	AAH49504	Human GTP-binding

ALIGNMENTS

RESULT 1
AAZ12026
ID AAZ12026 standard; DNA; 1344 BP.
XX
AC AAZ12026;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis complete ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38561.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122-123; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX

SQ Sequence 1344 BP; 329 A; 366 C; 365 G; 284 T; 0 other;

Query Match 100.0%; Score 1344; DB 20; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540

Qy    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660

Qy    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720

Qy    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344
```

```
RESULT   2
AAA81495/c
ID   AAA81495 standard; DNA; 44608 BP.
XX
AC   AAA81495;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_42 SEQ ID NO:42.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1283-1296; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 44608 BP; 10938 A; 10835 C; 11999 G; 10834 T; 2 other;
```

```
  Query Match             100.0%;  Score 1344;  DB 21;  Length 44608;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25944 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 25885

Qy      61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25884 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 25825

Qy     121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25824 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 25765

Qy     181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25764 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 25705

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25704 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 25645

Qy     301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25644 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 25585

Qy     361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25584 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 25525

Qy     421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25524 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 25465

Qy     481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25464 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 25405

Qy     541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25404 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 25345

Qy     601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25344 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 25285

Qy     661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25284 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 25225
```

```
Qy    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25224 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 25165

Qy    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25164 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 25105

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25104 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 25045

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  25044 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 24985

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24984 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 24925

Qy   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24924 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 24865

Qy   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24864 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 24805

Qy   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24804 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 24745

Qy   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24744 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 24685

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  24684 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 24625

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db  24624 GAAACCATTGAGAAGGAAGGCTGA 24601
```

RESULT 3
AAF21607/c
ID AAF21607 standard; DNA; 349980 BP.
XX
AC AAF21607;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:108.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 349980 BP; 84410 A; 84863 C; 94187 G; 86520 T; 0 other;

Query Match 100.0%; Score 1344; DB 21; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297909 ATGATTAAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGGCAGACCGGAGCAAGCCGTT 297850

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297849 TACGACGGCCCGGCCATTACCGAAGTCGCGTTGCTTGGCGAAGAATATGCCGGTATGCGC 297790

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297789 CCCTCGATGAAAGTCAAGGAAGGCGATGCCGTCAAAAAAGGCCAAGTGCTGTTTGAAGAC 297730

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297729 AAAAAGAATCCGGGCGTGGTGTTTACTGCGCCGGCTTCAGGCAAAATCGCCGCGATTCAC 297670

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297669 CGTGGCGAAAAGCGCGTACTTCAGTCAGTCGTGATTGCCGTTGAAGGCAACGACGAAATC 297610

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297609 GAGTTTGAACGCTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGC 297550

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297549 AACCTGATCCAATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCT 297490

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 297489 GCCGTCGATGCCGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCT 297430

Qy    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297429 GCCGACCCTACGGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTA 297370

Qy    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297369 TTGAGCCGTTTGACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCG 297310

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297309 TCTGAAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTG 297250

Qy    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297249 AGTGGCACGCACATTCATTTCATCGAGCCGGTCGGCGCGAATAAAACCGTGTGGACCATC 297190

Qy    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297189 AATTATCAAGATGTAATTACCATTGGCCGTTTGTTTGCAACAGGCCGTCTGAACACCGAG 297130

Qy    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297129 CGCGTGATTGCCCTAGGTGGTTCTCAAGTCAACAAACCGCGCCTCTTGCGTACCGTTTTG 297070

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297069 GGTGCGAAAGTATCGCAAATTACTGCGGGCGAATTGGTTGACACAGACAACCGCGTGATT 297010

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 297009 TCCGGTTCGGTATTGAACGGCGCGATTACACAAGGCGCGCACGATTATTTGGGACGCTAC 296950

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296949 CACAATCAGATTTCCGTTATCGAAGAAGGCCGCAGCAAAGAGCTGTTCGGCTGGGTTGCG 296890

Qy   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296889 CCGCAGCCGGACAAATACTCCATCACGCGTACAACCCTCGGCCATTTCCTGAAAAACAAA 296830

Qy   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296829 CTCTTCAAGTTCAACACAGCCGTCAACGGCGGCGACCGCGCCATGGTGCCGATTGGTACT 296770

Qy   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296769 TACGAGCGCGTGATGCCCTTGGATATCCTGCCCACCCTGCTTTTGCGCGATTTAATCGTC 296710

Qy   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296709 GGCGATACCGACAGCGCGCAGGCATTGGGTTGCTTGGAATTGGACGAAGAAGACCTCGCT 296650

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296649 TTGTGCAGCTTCGTCTGCCCGGGCAAATACGAATACGGCCCGCTGTTGCGCAAAGTGCTG 296590

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db 296589 GAAACCATTGAGAAGGAAGGCTGA 296566


RESULT   4
AAZ12027
ID   AAZ12027 standard; DNA; 1344 BP.
XX
AC   AAZ12027;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria meningitidis strain A complete ORF22 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38562.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 123; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 1344 BP; 323 A; 347 C; 368 G; 289 T; 17 other;


  Query Match             91.5%;  Score 1229.4;  DB 20;  Length 1344;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches 1266;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagtcatt 60

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          || ||||| || | ||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tatgacgggcccgtcattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccctngatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          |||||| |||||||||||||||||| ||||| | |||||||||||||||||| || ||
Db    181 aaaaagnatccgggcgtggtgtttaccgcgccngtttcaggcaaaatcgccgccatccat 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          || ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    241 cgcggcgaaaagcgcgtacttcagtcggtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||| |||||||||||| || |||||| |||||||||||||||||| ||| |  |  ||
Db    301 gagttcgaacgctacgcgcccgaagcgttggcaaacttaagcggcgangaaantnngnngc 360

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          || |||||||||||||||||||||||||||||||| | |||||||||||||||||| |||
```

```
Db    361 aatctgatccaatccggtttgtggactgcgctgcgtanccgtccgttcagcaaaatccct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctngcg 480

Qy    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540
          || ||||||  |||  | ||||||||||||| ||| ||||||| |||  |  || ||||||
Db    481 gcagaccctgtggttgtgatcaaagaagccgncgangatttcagacgangtntgctggta 540

Qy    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600
          |||||||||||||||| || |||||||||||| ||||||||||||||||||||||||||||
Db    541 ttgagccgtttgaccgagcgtaaaatccatgtgtgtaaggcagctggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatccggccggtttg 660

Qy    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720
          |||||||||||||||||||||| |||||||||| || || |||||||| |||||||||
Db    661 agtggcacgcacattcatttcattgagccggtcggtgcaaacaaaaccgtttggaccatc 720

Qy    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780
          |||||||||||||||| |||| || ||||||||||||||||||||||||||||||||||||
Db    721 aattatcaagatgtaattgccatcggacgtttgtttgcaacaggccgtctgaacaccgag 780

Qy    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          |||||||||||  | |||||||||||||||||||||| ||||||||||||||||||||||
Db    781 cgcgtgattgctttgggtggttctcaagtcaacaaaccacgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacgcagacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgtacgaccctcggccatttcctgaaaaacaaa 1080

Qy   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||||  |||||||||||||| ||||||||||||||||||||||||||||||
Db   1081 ctcttcaagttcacgacagccgtcaacggtggcgaccgcgccatggtgccgattggtact 1140

Qy   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200
          |||||||||||  |||||  | || |||||||| ||||||||||||||||||||||||||||
Db   1141 tacgagcgcgtaatgccgctagacatcctgcctaccctgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaagcattgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          |||||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatanggcccgctgttgcgtaaggtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||| ||||||||||||||||||
Db   1321 gaaaccnttgagaaggaaggctga 1344


RESULT   5
AAZ12028
ID   AAZ12028 standard; DNA; 1344 BP.
XX
AC   AAZ12028;
XX
```

```
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria gonorrhoeae  ORF22 polynucleotide sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38565.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 125; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 1344 BP; 334 A; 363 C; 362 G; 285 T; 0 other;


  Query Match             90.2%;  Score 1212.8;  DB 20;  Length 1344;
  Best Local Similarity   93.9%;  Pred. No. 0;
  Matches 1262;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||| | ||
Db      1 atgattaaaatcaaaaaaggtctaaatctgcccatcgcgggcagaccggagcaagtcatt 60

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          || |||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db     61 tatgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgtcggcatgcgc 120

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          |||||||||||| |||||||| || ||||||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaaatcaaggaaggtgaagccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          |||||||||||||||| || |||||||||||||||||||||||||||||| ||||||
Db    181 aaaaagaatccgggcgtagtatttactgcgccggcttcaggcaaaatcgccgctattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||| |||||||||| |||||||||||||||||| || ||| ||||| ||||||||||
```

```
Db    301 gagttcgaacgctacgtacctgaagcgctggcaaaattgagcagcgaaaaagtgcgccgc 360

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          |||||||| ||||| || || |||||||||| |||||||||||||||||||||| |||
Db    361 aacctgattcaatcaggcttatggactgcgcttcgcacccgtccgttcagcaaaatccct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtagatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Qy    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540
          |||||||||||||||| |||||||||||||||| || ||||||||||||||||||||
Db    481 gccgaccctacggtcatcatcaaagaagccgccgaagacttcaaacgcggcctgttggta 540

Qy    541 ttgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccg 600
          |||||||| |||||||||| |||||||||| ||||| ||||| ||||||||||||||||
Db    541 ttgagccgcctgaccgaacgtaaaatccatgtgtgtaaagcagcaggcgcagacgtgccg 600

Qy    601 tctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttg 660
          |||||||||||||||| |||||||||||||||| |||||||||||||||||||| |||
Db    601 tctgaaaatgctgccaatatcgaaacacatgaatttggcggcccgcatcctgccggcttg 660

Qy    661 agtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggaccatc 720
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    661 agtggcacgcacattcatttcatcgagccagtcggcgcgaataaaaccgtgtggaccatc 720

Qy    721 aattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccgag 780
          |||||||||| || ||| | || || |||||||| | ||||||||||||||| ||||||
Db    721 aattatcaagacgtgattgctatcggacgtttgttcgtaacaggccgtctgaatacccgag 780

Qy    781 cgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840
          |||||| ||||| | || ||    ||||||||||||||||||||||||||||||||||||
Db    781 cgcgtggttgccttgggcggcctgcaagtcaacaaaccgcgcctcttgcgtaccgttttg 840

Qy    841 ggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtgatt 900
          |||||||| || || ||| |||| || ||||||||||||||| | ||||||||||||||
Db    841 ggtgcgaaggtgtctcaacttaccgccggcgaattggttgacgcggacaaccgcgtgatt 900

Qy    901 tccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgctac 960
          |||||||||||||||||||| |||||| |||||||||||||| ||||||||||||||||||
Db    901 tccggttcggtattgaacggtgcgattgcacaaggcgcgcatgattatttgggacgctac 960

Qy    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 cacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcg 1020

Qy   1021 ccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaa 1080
          |||||||||||||||||||||||||||||| || || |||||||||||||| |||||||||
Db   1021 ccgcagccggacaaatactccatcacgcgcaccactctcggccatttcctaaaaaacaaa 1080

Qy   1081 ctcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggtact 1140
          |||||||||||||  |||||||||||||||||||||||||||||||| ||||| || |||
Db   1081 ctcttcaagttcacgacagccgtcaacggcggcgaccgcgccatggtaccgatcggcact 1140

Qy   1141 tacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcgtc 1200
          || |||||||| ||||| ||||||||| ||| ||||||||||||||||||||||||||||
Db   1141 tatgagcgcgtaatgccgttggacatcctgcctaccttgcttttgcgcgatttaatcgtc 1200

Qy   1201 ggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcgct 1260
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   1201 ggcgataccgacagcgcgcaggctttgggttgcttggaattggacgaagaagacctcgct 1260

Qy   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctg 1320

Qy   1321 gaaaccattgagaaggaaggctga 1344
          |||||||||||||||||||||||
Db   1321 gaaaccattgagaaggaaggctga 1344
```

RESULT 6
AAT42063

```
ID   AAT42063 standard; DNA; 1830121 BP.
XX
AC   AAT42063;
XX
DT   14-SEP-1999  (first entry)
XX
DE   Haemophilus influenzae complete genome sequence.
XX
KW   Genome; bacterium; Haemophilus influenzae; computer readable medium;
KW   expression modulating fragment; regulation; gene expression; vector;
KW   organism; open reading frame; ORF; ds.
XX
OS   Haemophilus influenzae.
XX
PN   WO9633276-A1.
XX
PD   24-OCT-1996.
XX
PF   22-APR-1996;   96WO-US05320.
XX
PR   07-JUN-1995;   95US-0487429.
PR   21-APR-1995;   95US-0426787.
PR   07-JUN-1995;   95US-0476102.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Adams MD,  Fleischmann RD,  Smith HO,  Venter JC,  White O;
XX
DR   WPI; 1996-485782/48.
XX
PT   Haemophilus influenzae Rd genome recorded on computer readable
PT   medium  - useful for identifying commercially important nucleic acid
PT   fragments by homology searching
XX
PS   Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
CC   This sequence represents the complete genome sequence of the bacterium
CC   Haemophilus influenzae strain Rd.  The invention relates to a computer
CC   readable medium (CRM) having recorded upon it the complete H.influenzae
CC   nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC   sequence at least 99% identical to (I).  By providing the full-length
CC   genomic sequence in a computer readable form, it is possible to identify
CC   commercially important nucleic acid fragments and expression modulating
CC   fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC   regulate the expression of a nucleic acid molecule.  Vectors and altered
CC   organisms comprising the predicted ORFS can be used to produce any of the
CC   polypeptide fragments of the H. influenzae Rd genome.
XX
SQ   Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
```

```
  Query Match             38.8%;  Score 521.8;  DB 17;  Length 1830121;
  Best Local Similarity   62.7%;  Pred. No. 1.6e-147;
  Matches  845;  Conservative    0;  Mismatches  497;  Indels    5;  Gaps    2;

Qy       1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
           ||||||| ||| || |||||| |  | || || || ||||| | ||| |  ||||   |
Db  179007 atgattacaattaagaaaggtttggatctcccaattgcgggaaaaccagcacaagtaatc 179066

Qy      61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
             |   ||||   ||  | |   |||| ||| | || || || || ||||  || |||||
Db  179067 catagcggcaacgctgtgaatcaagttgcgattctaggtgaggagtatgtggggatgcgt 179126

Qy     121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
           || || ||||| ||    ||||||||||  || |||||||| ||||| || |||||||||
Db  179127 ccttcaatgaaggtgcgcgaaggcgatgttgtgaaaaaaggtcaagtacttttgaagac 179186

Qy     181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
           ||||| ||||| || ||  | ||||| || || ||    || |  ||| | || ||  |
Db  179187 aaaaaaaatcctggtgtaatttttacagcccctgcaagcggtaccatcactgcaatcaat 179246

Qy     241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
```

```
Db 179247 cgtggcgaaaaacgtgtattacaatctgttgttattaatgtagaaggtgatgaaaaaatc 179306

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Db 179307 acttttgcaaaatatagcacagaacaattgaatacactttcttctgaacaagtaaaacaa 179366

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Db 179367 aacttgatagaatctggcttatggacggcattacgtactcgtccatttagtaaagttcca 179426

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggct 480

Db 179427 tccattgaaagtgaagcatcttctatttttgtaaatgcgatggataccaatccattagca 179486

Qy    481 gccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggta 540

Db 179487 gccgatccctctgtcgtattaaaagaatattcgcaagacttcactaacggtttaaccgta 179546

Qy    541 ttgagccgt---ttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtg 597

Db 179547 ttaagtcgtctattcccttcaaaaccattacacttatgtaaagcaggcgattctaatatt 179606

Qy    598 ccgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggt 657

Db 179607 ccaactgctgatcttgaaaatttacagattcatgattttactggtgttcatccagctggt 179666

Qy    658 ttgagtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggacc 717

Db 179667 cttgtagggacgcatattcactttattgatcctgtcggtattcaaaaaactgtatggcat 179726

Qy    718 atcaattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacacc 777

Db 179727 atcaattatcaagatgtgatcgctgtaggtaaattatttacaacaggcgaactttattca 179786

Qy    778 gagcgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgtt 837

Db 179787 gaacgagtaatttctcttgcaggtcctcaagtgaaagaacctcgtttagtacgtacaaca 179846

Qy    838 ttgggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagacaaccgcgtg 897

Db 179847 atcggtgctaatctttctcaattaacccaaaatgaattaagtgcaggtaaaaatcgtgta 179906

Qy    898 atttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgggacgc 957

Db 179907 atttctggctcagttctttgtggtcaaatagcaaaagattcacatgattatttaggtcgc 179966

Qy    958 taccacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttcggctgggtt 1017

Db 179967 tatgctttgcaagtatcagtaatcgctgaaggtaatgaaaaagaatttttggttggatt 180026

Qy   1018 gcgccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcctgaaaaac 1077

Db 180027 atgccacaagccaataaatattctgtgactcgtaccgtattagggccacttcagtaaa-- 180084

Qy   1078 aaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggt 1137

Db 180085 aaattgtttaattttactacttcagaaaatggtggcgaacgtgcaatggtgccaatcggt 180144

Qy   1138 acttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatc 1197

Db 180145 agctatgagcgcgtaatgccattggacattttacctacattattattacgtgatttaatc 180204

Qy   1198 gtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctc 1257

Db 180205 gtgggcgataccgatggtgctcaagagttaggttgtctggaattagacgaagaagacttg 180264

Qy   1258 gctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtg 1317

Db 180265 gcgttatgttctttcgtttgtccaggcaaatatgaatacggttcaatcttgcgtcaagtc 180324

Qy   1318 ctggaaaccattgagaaggaaggctga 1344
```

Db 180325 ttagataagattgagaaggaaggttaa 180351

RESULT 7
AAZ12025
ID AAZ12025 standard; DNA; 474 BP.
XX
AC AAZ12025;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis partial ORF22 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38560.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 122; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;

Query Match 35.2%; Score 473.6; DB 20; Length 474;
Best Local Similarity 99.8%; Pred. No. 1.7e-134;
Matches 473; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
```

```
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
```

RESULT 8
AAA81335
ID AAA81335 standard; DNA; 474 BP.
XX
AC AAA81335;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis MenB polynucleotide sequence ORF number 14.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Disclosure; Page 203; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;

```
  Query Match             35.2%;  Score 473.6;  DB 21;  Length 474;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-134;
  Matches  473;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaatc 300
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaargcaacgacgaaatc 300

Qy    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgc 360

Qy    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 aacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcct 420

Qy    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
```

RESULT 9
AAZ54034/c
ID AAZ54034 standard; DNA; 363 BP.
XX
AC AAZ54034;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2017.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.

```
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75272.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 1003; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 74 A; 96 C; 110 G; 83 T; 0 other;

  Query Match             27.0%;  Score 363;  DB 21;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-101;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     312 ctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgcaacctgatcca 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 CTACGCACCTGAAGCGCTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGCAACCTGATCCA 304

Qy     372 atccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcctgccgtcgatgc 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATCCGGTTTGTGGACTGCGCTGCGCACCCGTCCGTTCAGCAAAATTCCTGCCGTCGATGC 244

Qy     432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy     492 ggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggtattgagccgttt 551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 GGTCATTATCAAAGAAGCCGCCGAGGATTTCAAACGCGGCCTGTTGGTATTGAGCCGTTT 124

Qy     552 gaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 GACCGAACGCAAAATCCATGTTTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy     612 tgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttgagtggcacgca 671
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCTGCCGGTTTGAGTGGCACGCA 4

Qy     672 cat 674
           |||
Db       3 CAT 1
```

```
RESULT  10
AAZ54035/c
ID   AAZ54035 standard; DNA; 363 BP.
XX
AC   AAZ54035;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 628 partial DNA sequence SEQ ID NO:2019.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75273.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 1003-1004; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 363 BP; 75 A; 105 C; 103 G; 80 T; 0 other;

  Query Match             24.4%;  Score 327.8;  DB 21;  Length 363;
  Best Local Similarity   93.9%;  Pred. No. 5.7e-90;
  Matches  341;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     312 ctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgcaacctgatcca 371
           |||||| || |||||| ||||||||||||||||||||||||||||||||| ||||||||
Db     363 CTACGCGCCCGAAGCGTTGGCAAACTTAAGCGGCGAAGAAGTGCGCCGCAATCTGATCCA 304
```

```
Qy    372 atccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcctgccgtcgatgc 431
          ||||||||||||||||||||||||| ||||||||||||||||||| ||||||||||||||
Db    303 ATCCGGTTTGTGGACTGCGCTGCGTACCCGTCCGTTCAGCAAAATCCCTGCCGTCGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
          |||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCGGCAGACCCTGT 184

Qy    492 ggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggtattgagccgttt 551
          |||  | ||||||||||||||||||||||||| ||| ||  || ||||||||||||||||
Db    183 GGTTGTGATCAAAGAAGCCGCCGAGGATTTCAGACGAGGTTTGCTGGTATTGAGCCGTTT 124

Qy    552 gaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
          |||||| || |||||||||| ||||||||||||||||||||||||||||||||||||||
Db    123 GACCGAGCGTAAAATCCATGTGTGTAAGGCAGCTGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttgagtggcacgca 671
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     63 TGCCAACATCGAAACACATGAATTCGGCGGCCCGCATCCGGCCGGTTTGAGTGGCACGCA 4

Qy    672 cat 674
          |||
Db      3 CAT 1
```

RESULT 11
AAF28548/c
ID AAF28548 standard; DNA; 96109 BP.
XX
AC AAF28548;
XX
DT 04-APR-2001 (first entry)
XX
DE Genomic fragment #35.
XX
KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW bronchopulmonary; endocarditis; meningitis; ss.
XX
OS Moraxella catarrhalis.
XX
PN WO200078968-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16649.
XX
PR 18-JUN-1999; 99US-0140121.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lagace RE, Patterson C, Berg KL;
XX
DR WPI; 2001-041427/05.
XX
PT Genomic library for identifying diagnostic and therapeutic
PT compositions, and for identifying virulence factors, regulatory
PT elements and drug targets, comprises Moraxella catarrhalis nucleic
PT acids -
XX
PS Claim 1; Page 345-368; 545pp; English.
XX
CC The present invention relates to a Moraxella catarrhalis genomic library
CC comprising of a combination of 41 nucleic acid molecules (see
CC AAF28514-AAF28554). The library has a number of uses described in the
CC specification e.g. is useful for aidentifying diagnostic and therapeutic
CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC aerobic, gram-negative diplococcus, normally found among the bacterial
CC flora of human upper airways. M. catarrhalis is known to cause acute,
CC localised infections such as otitis media, sinusitis and bronchopulmonary
CC infection and life-threatening, systemic diseases including endocarditis
CC and meningitis.
XX
SQ Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 other;

Query Match 24.3%; Score 326.4; DB 22; Length 96109;
Best Local Similarity 54.2%; Pred. No. 2.3e-88;
Matches 735; Conservative 0; Mismatches 606; Indels 15; Gaps 3;

```
Qy      1 atgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtt 60
          |||||||  ||||||||||||  |  | |||||  ||| | |||   ||     |     ||
Db  85201 ATGATTACCATCAAAAAAGGCTTGGATCTGCCAATCACTGGCGAGCCCTCCAAGCAGATT 85142

Qy     61 tacgacggcccggccattaccgaagtcgcgttgcttggcgaagaatatgccggtatgcgc 120
            |||   | ||  |||  ||   | ||  |||||  | |||  | ||  |||| |||  |||||
Db  85141 AGCGAGCACACGCCCACTA---AGGTGGCGTTAGTCGGCTACGATTATGTCGGCATGCGT 85085

Qy    121 ccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagac 180
          ||  | |||  | |||||  |||||  |||   ||||  |||||  ||  |||||  |  ||||||
Db  85084 CCGACCATGCATGTCAAAGAAGGTGATAAGGTCACCAAAGGACAGGTGCTTTACGAAGAC 85025

Qy    181 aaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcac 240
          |||||     | |||||  | | |||  ||  ||  | |   ||||   |||  ||  |  |
Db  85024 AAAAAACGAGCAGGCGTCCTTTATACAGCACCAGTTGGCGGCACTGTCGTTGCTGTCAAT 84965

Qy    241 cgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttg------aaggcaacgac 294
          |||||||||    || ||| || |      | || ||  |||||  |       |  |  ||
Db  84964 CGTGGCGAACGACGAGTATTTGAAAGCTTGGTCATCGCCGTCGATCCAAATGCAAGTGAA 84905

Qy    295 gaaatcgagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtg 354
          ||||||   ||| |    ||     ||||||   || |||   |     || |  |||
Db  84904 GAAATCACTTTTAATGCTTATCGCTCTGAAGATTTGTCAATGCTGGATCGCAAGACAGTT 84845

Qy    355 cgccgcaacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaa 414
                |  || ||   || |||   ||||| || || |||||  ||||| || |||
Db  84844 GTTGAGCAGTTGGTCGCTTCAGGTGAATGGACCGCCCTACGCACTCGTCCTTTTAGCCGT 84785

Qy    415 attcctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccg 474
          |  || |  |  ||  |   || |  ||||| || || |  || |  || ||||||||
Db  84784 ACACCAGAAATTAATAGCATACCTTCTGCCATTTTTGTGACGGCAACAGATACCAATCCT 84725

Qy    475 ctggctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctg 534
          ||||||  ||||||  |   | |  |    ||||      | |  || ||    ||| ||
Db  84724 TTGGCTGTTGACCCTGCCACCGTATTATTTGAAGAAATACAAGCATTTAATGATGGCTTG 84665

Qy    535 ttggtattgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagac 594
            | ||         |  ||  |  |          | |  |   ||  | ||
Db  84664 GCAATTTTATCAACGCTATCCCCAAAAACCTATGTCTGTCATGGTGATATTGTACCTGAG 84605

Qy    595 gtgccgtctgaaaatgctgccaacatcgaaacac------atgaattcggcggcccgcat 648
                | |||  |||| |   | ||| ||         | |  ||    ||    |||
Db  84604 ACCTTAAGTAAAACCGCTGACGGTAACGATACCATCTATGAAGGGTTTACTGGTGTACAT 84545

Qy    649 cctgccggtttgagtggcacgcacattcatttcatcgagccggtcggcgcgaataaaacc 708
          || || ||       || |||||||||| ||  |  | ||  | |        |
Db  84544 CCAGCAGGCAACGCAGGTACGCACATTCACTTTTTACACCCTTTGGCACGAGGCAGCTTA 84485

Qy    709 gtgtggaccatcaattatcaagatgtaattaccattggccgtttgtttgcaacaggccgt 768
          || ||||| || ||||||||||||||| ||  |||||||     | ||| | ||||| ||
Db  84484 GTTTGGACGATTAATTATCAAGATGTCATCGCCATTGGTAAGCTATTTACCACAGGTCGC 84425

Qy    769 ctgaacaccgagcgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttg 828
          |  |||| |||||  ||||     | |  ||| |   |||   ||| || ||| |  ||
Db  84424 ATTTACACAGAGCGTCTGATCAGTTTGGCAGGTCCAGCAGTGGCCAATCCTCGCTTGGTG 84365

Qy    829 cgtaccgttttgggtgcgaaagtatcgcaaattactgcgggcgaattggttgacacagac 888
          ||||||  |  ||||  |  | |  || | ||    || ||| |    | |  ||
Db  84364 CGTACCACTCGAGGTGCTGATTTGACCGAATTGACCCAAGGTGAACTCAACGGCGACGAT 84305

Qy    889 aaccgcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattat 948
          ||||| || ||||| || || ||  |     ||     |    |    || ||    | | |
Db  84304 AACCGTGTCATTTCAGGCTCTGTTCTATCAGGTCGCACAGCTGTTGGTGCTACGGCATTT 84245

Qy    949 ttgggacgctaccacaatcagatttccgttatcgaagaaggccgcagcaaagagctgttc 1008
          ||||| ||||  || |||||  | || ||  |    |||||  ||                |
Db  84244 TTGGGCCGCTTTCATAATCAAGTCTCTGTACTTACCGAAGGTCGTGAACGCCCTGCACTT 84185
```

```
Qy   1009 ggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccatttc 1068
Db  84184 CACTTTTTTACACCAGGGGCAAATCGTTTTTCAAAATTGCCAATTTATATTTCACAGTTT 84125

Qy   1069 ctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtg 1128
Db  84124 TTTAAAGGCAAAAAATACGATTTTACGACCAGTACCAACGGCTCGCCTCGTGCCATGGTA 84065

Qy   1129 ccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgc 1188
Db  84064 CCAATCGGCTTATTTGAAACTGTCATGCCTCAAGACTATCTACCTACCCAGCTACTGCGT 84005

Qy   1189 gatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaa 1248
Db  84004 GCTTTGATTGTCGAAGATATCATCAGTGCGGTTGAATTGGGTGCACTTGAGCTTGATGAA 83945

Qy   1249 gaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgttg 1308
Db  83944 GAAGATTTGGCACTGTGTACTTTCGTATCGCCAGGCAAATATGAGTTCGGTGATATTTTG 83885

Qy   1309 cgcaaagtgctggaaaccattgagaaggaaggctga 1344
Db  83884 CGTGATAACTTGACCCGTATTGAGCAGGAAGGCTAA 83849
```

RESULT 12
AAZ54033/c
ID AAZ54033 standard; DNA; 363 BP.
XX
AC AAZ54033;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrheae ORF 628 partial DNA sequence SEQ ID NO:2015.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria gonorrheae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; AAY75271.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 7; Page 1003; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 363 BP; 71 A; 92 C; 109 G; 91 T; 0 other;

Query Match 24.2%; Score 324.6; DB 21; Length 363;
Best Local Similarity 93.4%; Pred. No. 5.5e-89;
Matches 339; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy    312 ctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgccgcaacctgatcca 371
Db    363 CTACGTACCTGAAGCGCTGGCAAAATTGAGCAGCGAAAAAGTGCGCCGCAACCTGATTCA 304

Qy    372 atccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattcctgccgtcgatgc 431
Db    303 ATCAGGCTTATGGACTGCGCTTCGCACCCGTCCGTTCAGCAAAATCCCTGCCGTAGATGC 244

Qy    432 cgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctggctgccgaccctac 491
Db    243 CGAGCCGTTCGCCATCTTCGTCAATGCGATGGACACCAATCCGCTGGCTGCCGACCCTAC 184

Qy    492 ggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttggtattgagccgttt 551
Db    183 GGTCATCATCAAAGAAGCCGCCGAAGACTTCAAACGCGGCCTGTTGGTATTGAGCCGCCT 124

Qy    552 gaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgccgtctgaaaatgc 611
Db    123 GACCGAACGTAAAATCCATGTGTGTAAAGCAGCAGGCGCAGACGTGCCGTCTGAAAATGC 64

Qy    612 tgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtttgagtggcacgca 671
Db     63 TGCCAATATCGAAACACATGAATTTGGCGGCCCGCATCCTGCCGGCTTGAGTGGCACGCA 4

Qy    672 cat 674
Db      3 CAT 1
```

RESULT 13
AAX91657
ID AAX91657 standard; DNA; 1353 BP.
XX
AC AAX91657;
XX
DT 25-AUG-1999 (first entry)
XX
DE Porphorymonas gingivalis protein PG1 encoding DNA.
XX
KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic; ds.
XX
OS Porphorymonas gingivalis.
XX
PN WO9929870-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-AU01023.
XX
PR 04-AUG-1998; 98AU-0005028.
PR 10-DEC-1997; 97AU-0000839.
PR 31-DEC-1997; 97AU-0001182.

```
PR   30-JAN-1998;   98AU-0001546.
PR   10-MAR-1998;   98AU-0002264.
PR   09-APR-1998;   98AU-0002911.
PR   23-APR-1998;   98AU-0003128.
PR   05-MAY-1998;   98AU-0003338.
PR   22-MAY-1998;   98AU-0003654.
PR   29-JUL-1998;   98AU-0004917.
XX
PA   (CSLC-) CSL LTD.
XX
PI   Agius CT,  Barr IG,  Hocking DM,  Margetts MB,  Patterson MA;
PI   Ross BC,  Rothel LJ,  Webb EA;
XX
DR   WPI; 1999-385613/32.
DR   P-PSDB; AAY34439.
XX
PT   Antigenic Porphorymonas gingivalis peptides for preventing
PT   gingivitis
XX
PS   Claim 12; Page 179; 588pp; English.
XX
CC   AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC   Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC   AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC   isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC   activity with a vaccine mechanism of action. The PG polypeptides can be
CC   used as vaccines especially against Porphorymonas gingivalis. Probes can
CC   be used to detect Porphorymonas gingivalis in standard hybridisation
CC   assays. Porphorymonas gingivalis is involved in periodontal disease
CC   especially gingivitis.
XX
SQ   Sequence 1353 BP; 337 A; 333 C; 379 G; 304 T; 0 other;
```

Query Match 8.5%; Score 113.6; DB 20; Length 1353;
Best Local Similarity 46.1%; Pred. No. 3.5e-24;
Matches 623; Conservative 0; Mismatches 699; Indels 30; Gaps 6;

```
Qy      2 tgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtttt 61
Db     11 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 70

Qy     62 acgacggcccggcc---attaccgaagtcgcgttgcttggcgaagaatatgccggtatgc 118
Db     71 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 130

Qy    119 gcccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaag 178
Db    131 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 190

Qy    179 acaaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattc 238
Db    191 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 250

Qy    239 accgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaa 298
Db    251 atcgcggtgccaagcgcaaggtgttgagcatcgaggtgaaaccggacggactgaacgaat 310

Qy    299 tcgagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgcc 358
Db    311 acgagt------cattccctgtcggggatccgtctgccctctctgccgaacagatcaagg 364

Qy    359 gcaacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattc 418
Db    365 agcttttactgtcgagcggtatgtggggttttattaagcaacgtccttacgacatagtgg 424

Qy    419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
Db    425 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 484

Qy    479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttgg 538
Db    485 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 544

Qy    539 tattgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgc 598
Db    545 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 604

Qy    599 cgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtt 658
Db    605 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 655

Qy    659 tgagtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggacca 718
Db    656 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 715

Qy    719 tcaattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccg 778
Db    716 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 775

Qy    779 agcgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttt 838
Db    776 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 835

Qy    839 tgggtgcgaaagtatcgcaaattactgcgggcgaattg------gttgacacagacaacc 892
Db    836 cgggttgcaatgtctttgcttccttccccggccgactgacaataaaggaatctcacgagc 895

Qy    893 gcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgg 952
Db    896 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagcctttcctgt 955

Qy    953 gacgctaccacaatcagatttccgttatcgaagaaggccgc---agcaaagagctgttcg 1009
Db    956 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1015

Qy   1010 gctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcc 1069
Db   1016 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1075

Qy   1070 ---tgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatgg 1126
Db   1076 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1135

Qy   1127 tgccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgc 1186
Db   1136 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1195

Qy   1187 gcgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacg 1246
Db   1196 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1255

Qy   1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgt 1306
Db   1256 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1315

Qy   1307 tgcgcaaagtgctggaaaccattgagaaggaa 1338
Db   1316 ttcgcgagggcttggatatgctctataaggaa 1347
```

```
RESULT  14
AAX91536
ID   AAX91536 standard; DNA; 1362 BP.
XX
AC   AAX91536;
XX
DT   25-AUG-1999  (first entry)
XX
DE   Porphorymonas gingivalis protein PG1 ORF encoding DNA.
XX
KW   Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW   vaccine; antigenic; ds.
XX
OS   Porphorymonas gingivalis.
XX
```

```
PN   WO9929870-A1.
XX
PD   17-JUN-1999.
XX
PF   10-DEC-1998;   98WO-AU01023.
XX
PR   04-AUG-1998;   98AU-0005028.
PR   10-DEC-1997;   97AU-0000839.
PR   31-DEC-1997;   97AU-0001182.
PR   30-JAN-1998;   98AU-0001546.
PR   10-MAR-1998;   98AU-0002264.
PR   09-APR-1998;   98AU-0002911.
PR   23-APR-1998;   98AU-0003128.
PR   05-MAY-1998;   98AU-0003338.
PR   22-MAY-1998;   98AU-0003654.
PR   29-JUL-1998;   98AU-0004917.
XX
PA   (CSLC-) CSL LTD.
XX
PI   Agius CT,  Barr IG,  Hocking DM,  Margetts MB,  Patterson MA;
PI   Ross BC,  Rothel LJ,  Webb EA;
XX
DR   WPI; 1999-385613/32.
DR   P-PSDB; AAY34318.
XX
PT   Antigenic Porphorymonas gingivalis peptides for preventing
PT   gingivitis
XX
PS   Claim 12; Page 96; 588pp; English.
XX
CC   AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC   Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC   AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC   isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC   activity with a vaccine mechanism of action. The PG polypeptides can be
CC   used as vaccines especially against Porphorymonas gingivalis. Probes can
CC   be used to detect Porphorymonas gingivalis in standard hybridisation
CC   assays. Porphorymonas gingivalis is involved in periodontal disease
CC   especially gingivitis.
XX
SQ   Sequence 1362 BP; 337 A; 335 C; 381 G; 309 T; 0 other;
```

```
  Query Match             8.5%;  Score 113.6;  DB 20;  Length 1362;
  Best Local Similarity   46.1%;  Pred. No. 3.5e-24;
  Matches  623;  Conservative    0;  Mismatches  699;  Indels   30;  Gaps    6;

Qy       2 tgattaaaatcaaaaaaggtctaaacctgcccatcgcgggcagaccggagcaagccgtttt 61
Db      20 ttataaaaacaaaaaaaggccttgcacttaatctgaaaggaaaaccgctgcccgagatgc 79

Qy      62 acgacggcccggcc---attaccgaagtcgcgttgcttggcgaagaatatgccggtatgc 118
Db      80 tggccgaaccggcccaaagtcctacttacgcggtcgtgcccgacgattttgaaggtgtta 139

Qy     119 gcccctcgatgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaag 178
Db     140 tccccaaggtgacggctcgtccgggggataaggtgcgtgccggctcagcactgatgcacc 199

Qy     179 acaaaaagaatccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattc 238
Db     200 acaaggcatatccggagatgaagtttacaagtccggttagcggcgaagtgatcgcggtga 259

Qy     239 accgtggcgaaaagcgcgtacttcagtcagtcgtgattgccgttgaaggcaacgacgaaa 298
Db     260 atcgcggtgccaagcgcaaggtgttgagcatcgaggtgaaaccggacggactgaacgaat 319

Qy     299 tcgagtttgaacgctacgcacctgaagcgctggcaaacttaagcggcgaagaagtgcgcc 358
Db     320 acgagt------cattccctgtcggggatccgtctgccctctctgccgaacagatcaagg 373

Qy     359 gcaacctgatccaatccggtttgtggactgcgctgcgcacccgtccgttcagcaaaattc 418
```

```
Db     374 agcttttactgtcgagcggtatgtggggttttattaagcaacgtccttacgacatagtgg 433

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
Db     434 ctacaccggatatagctccacgcgacatttatattactgccaactttactgcaccattgg 493

Qy     479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggcctgttgg 538
Db     494 ctccggacttcgatttcatcgttcgaggagaagaacgcgccctgcagactgccatcgatg 553

Qy     539 tattgagccgtttgaccgaacgcaaaatccatgtttgtaaggcagctggcgcagacgtgc 598
Db     554 ccttggccaaactcacgacaggaaaggtgtatgtgggcctgaagccgggttcatctctgg 613

Qy     599 cgtctgaaaatgctgccaacatcgaaacacatgaattcggcggcccgcatcctgccggtt 658
Db     614 gcttgcacaatgcagaaatcgtagaagtaca---------cggacctcatccggcaggta 664

Qy     659 tgagtggcacgcacattcatttcatcgagccggtcggcgcgaataaaaccgtgtggacca 718
Db     665 acgtgggcgtgctgatcaatcatacgaagccaatcaatcggggcgaaacggtgtggacgc 724

Qy     719 tcaattatcaagatgtaattaccattggccgtttgtttgcaacaggccgtctgaacaccg 778
Db     725 tcaaggctaccgacctgatcgtgatcggacgtttcctgcttacgggcaaagccgatttta 784

Qy     779 agcgcgtgattgccctaggtggttctcaagtcaacaaaccgcgcctcttgcgtaccgttt 838
Db     785 ccagaatgattgccatgaccggctcagacgctgcagctcacggatacgtccgtattatgc 844

Qy     839 tgggtgcgaaagtatcgcaaattactgcgggcgaattg------gttgacacagacaacc 892
Db     845 cgggttgcaatgtctttgcttccttccccggccgactgacaataaaggaatctcacgagc 904

Qy     893 gcgtgatttccggttcggtattgaacggcgcgattacacaaggcgcgcacgattatttgg 952
Db     905 gtgtgatcgatggcaatgtgctgaccggtaagaagctctgcgagaaggagccctttcctgt 964

Qy     953 gacgctaccacaatcagatttccgttatcgaagaaggccgc---agcaaagagctgttcg 1009
Db     965 cagcccggtgtgaccagatcacggtgatccccgaaggcgacgatgtggacgaactcttcg 1024

Qy    1010 gctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccatttcc 1069
Db    1025 ggtgggctgcaccccgtctcgatcagtacagcatgagcagagcttatttctcttggttgc 1084

Qy    1070 ---tgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatgg 1126
Db    1085 aggggaaaaacaaagagtacgtactcgatgcccggatcaagggtggcgaacgtgctatga 1144

Qy    1127 tgccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgc 1186
Db    1145 tcatgagcaacgagtatgaccgcgttttcccgatggacatctatccggagtatttgctca 1204

Qy    1187 gcgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacg 1246
Db    1205 aggctattatagcattcgacatcgacaagatggaggacttaggcatatatgaagtggctc 1264

Qy    1247 aagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgt 1306
Db    1265 cggaggactttgccacttgcgaatttgtggatacatccaagatcgagctgcagcgtatcg 1324

Qy    1307 tgcgcaaagtgctggaaaccattgagaaggaa 1338
Db    1325 ttcgcgagggcttggatatgctctataaggaa 1356
```

```
RESULT  15
AAX84332/c
ID   AAX84332 standard; DNA; 5059 BP.
XX
AC   AAX84332;
XX
```

```
DT   08-SEP-1999  (first entry)
XX
DE   Stealth virus nucleic acid clone, SEQ ID NO: 24.
XX
KW   Stealth virus; detection; diagnosis; infection; ss.
XX
OS   Stealth virus.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 3605
FT                   /*tag=  a
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3610
FT                   /*tag=  b
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3615
FT                   /*tag=  c
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3630
FT                   /*tag=  d
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3631
FT                   /*tag=  e
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3638
FT                   /*tag=  f
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3641
FT                   /*tag=  g
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3642
FT                   /*tag=  h
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3657
FT                   /*tag=  i
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3659
FT                   /*tag=  j
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3681
FT                   /*tag=  k
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3698
FT                   /*tag=  l
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3703
FT                   /*tag=  m
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3708
FT                   /*tag=  n
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3713
FT                   /*tag=  o
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3720
FT                   /*tag=  p
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3732
FT                   /*tag=  q
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3733
FT                   /*tag=  r
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3740
FT                   /*tag=  s
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3747
FT                   /*tag=  t
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3750
FT                   /*tag=  u
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3751
FT                   /*tag=  v
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3766
FT                   /*tag=  w
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3778
FT                   /*tag=  x
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3792
FT                   /*tag=  y
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3799
FT                   /*tag=  z
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3800
FT                   /*tag=  aa
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3812
FT                   /*tag=  ab
```

```
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3813
FT                   /*tag=  ac
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3818
FT                   /*tag=  ad
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3819
FT                   /*tag=  ae
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3824
FT                   /*tag=  af
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3827
FT                   /*tag=  ag
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3841
FT                   /*tag=  ah
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3853
FT                   /*tag=  ai
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3857
FT                   /*tag=  aj
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3860
FT                   /*tag=  ak
FT                   /note= "this nucleotide is represented as a * in the
FT                          specification, and is included to maintain the
FT                          base numbering given in the specification"
FT   misc_difference 3863
FT                   /*tag=  al

  Query Match             3.0%;  Score 40.8;  DB 20;  Length 5059;
  Best Local Similarity   16.7%;  Pred. No. 0.12;
  Matches   45;  Conservative  100;  Mismatches  125;  Indels    0;  Gaps    0;

Qy   1004 tgttcggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggcc 1063
          |::: : :|:: ::::|: | |:: : ::   :  |  :| |  :: |  |    :||
Db    598 TKYHGSYBTRVYDWSSGYAGNASHTSTHSYSTWNATATHTDATASSNAGNANATNYGGNT 539

Qy   1064 atttcctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgcca 1123
            |:  |      :::: |:  : : |::  : :|: :: :   :  |  : |   : :|
Db    538 GSTBAGTNCNNTKYRVHNTSNMYBHATYHNTSNMARGBSSYSNKKNAGHNSNCANTRAVA 479

Qy   1124 tggtgccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttt 1183
          |:| :  : ::    |: ::: : :  :| :|: |:: :  :| | :   :::  :  |:
Db    478 TKGASDNSSYDANATCYNWHSTSKSNTDGKDGSDCYDASSNHTGCNKKDAWYYANSTRTK 419

Qy   1184 tgcgcgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattgg 1243
           :::  |:   |:  :  : |: | || : : :::   |: ::  :::    | | :  |
Db    418 SSYBNTAYGGNAKGNKKNSAGHGANCGGSYBDSYBRGNGHNKKNTKDDAWNAGNANHGNG 359

Qy   1244 acgaagaagacctcgctttgtgcagcttcg 1273
          :  :|:: :|: | :  :|:|:   :|:::
```

```
Db    358 RTYDAKDGRAYATKKASYTSTKTYHYTKSK 329


RESULT  16
AAZ32020
ID   AAZ32020 standard; DNA; 38734 BP.
XX
AC   AAZ32020;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH1 related EST AL021529.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 296-321; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 38734 BP; 6142 A; 13140 C; 13585 G; 5867 T; 0 other;

  Query Match             2.8%;  Score 38.2;  DB 20;  Length 38734;
  Best Local Similarity   48.0%;  Pred. No. 2;
  Matches  109;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy   1116 ccgcgccatggtgccgattggtacttacgagcgcgtgatgcccttggatatcctgcccac 1175
           | ||| |||  |  || | |   |         ||||  |   | ||| | || | ||||
```

```
Db   4519 ctgcggcatcatcgcgctggccgccgcgaccgacgtgcggatgtcggagaacccggccac 4578

Qy   1176 cctgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgctt 1235
          |   || || | | |     |||  | ||   | |||  ||  |||     ||   ||
Db   4579 cgacctcttccacaacggcgtcgcggtcggcgcggactacgtccagcacccggtgatctc 4638

Qy   1236 ggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaata 1295
          | |  | | ||| |  | | |||    | |||||||| | | |   |       || |
Db   4639 gcaggtcgccgaggcggtcttcggcgagggcagcttcctgttcatcgtgctggccgcagc 4698

Qy   1296 cggcccgctgttgcgcaaagtgctggaaaccattgagaaggaaggct 1342
          |  | |||| | | |    |    |  | ||  | | |  | ||||
Db   4699 caccgcgctggtcctcttcctcgccgccaacaccgcgtacaacggct 4745


RESULT 17
AAC90077
ID   AAC90077 standard; DNA; 38734 BP.
XX
AC   AAC90077;
XX
DT   19-MAR-2001 (first entry)
XX
DE   AL021529 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;
PI   Fornwald JA, Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 7; Pages 597-622; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 38734 BP; 6142 A; 13138 C; 13586 G; 5868 T; 0 other;


  Query Match             2.8%;  Score 38.2;  DB 22;  Length 38734;
  Best Local Similarity   48.0%;  Pred. No. 2;
  Matches  109;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy   1116 ccgcgccatggtgccgattggtacttacgagcgcgtgatgcccttggatatcctgcccac 1175
          | ||| |||  |  || | |   |          ||||  |   | ||| | || | ||||
Db   4519 ctgcggcatcatcgcgctggccgccgcgaccgacgtgcggatgtcggagaacccggccac 4578

Qy   1176 cctgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgctt 1235
          |   || || | | |     |||  | ||   | |||  ||  |||     ||   ||
Db   4579 cgacctcttccacaacggcgtcgcggtcggcgcggactacgtccagcacccggtgatctc 4638

Qy   1236 ggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaata 1295
          | |  | | ||| |  | | |||    | |||||||| | | |   |       || |
Db   4639 gcaggtcgccgaggcggtcttcggcgagggcagcttcctgttcatcgtgctggccgcagc 4698

Qy   1296 cggcccgctgttgcgcaaagtgctggaaaccattgagaaggaaggct 1342
          |  | |||| | | |    |    |  | ||  | | |  | ||||
Db   4699 caccgcgctggtcctcttcctcgccgccaacaccgcgtacaacggct 4745


RESULT 18
AAS51578
ID   AAS51578 standard; DNA; 1416 BP.
XX
AC   AAS51578;
XX
DT   13-FEB-2002 (first entry)
XX
DE   Pseudomonas aeruginosa DNA for cellular proliferation protein #163.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI   Yamamoto RT, Xu HH;
XX
```

```
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33719.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4160; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1416 BP; 270 A; 522 C; 425 G; 199 T; 0 other;


  Query Match             2.8%;  Score 37.8;  DB 23;  Length 1416;
  Best Local Similarity   50.8%;  Pred. No. 0.55;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy     338 taagcggcgaagaagtgcgccgcaacctgatccaatccggtttgtggactgcgctgcgca 397
           | | ||||   || |||  |||| | | |||||    || ||   ||.  | ||||
Db     902 tcaccggcatcgacgtggtccgcgagcagatccgcatcgcttccgggctggagctgtcgg 961

Qy     398 cccgtccgttcagcaaaattcctgccgtcgatgccgagccgttcgccatcttcgtcaatg 457
             |   | |  |  ||   | | |  || | | |||  || ||||  ||||  || | | |
Db     962 tcaagcaggacgacatcgtccatcgcggctacgccctgcagttccgcatcaacgccgagg 1021

Qy     458 cgatggacaccaatccgctggctgccgaccctacggtcattatcaaagaagccgccg 514
               | | | |||    ||| |  |   |   | | |||    | |  | ||| |||
Db    1022 acccgaagaacaacttcctgccctcgttcggcaagatcacccgctactacgcccccg 1078


RESULT  19
AAC45663
ID   AAC45663 standard; DNA; 1539 BP.
XX
AC   AAC45663;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 47309.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.8%; Score 37.8; DB 21; Length 1539;
Best Local Similarity 55.8%; Pred. No. 0.57;
Matches 72; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy    128 tgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagacaaaaaga 187
          | ||||  |   |   |  ||   || ||||||    ||   ||||    |   |  |||
Db     49 tcaaagctacacagatctctgaaatcgaaaaagcttgagctgtgttaatggcggaggaga 108

Qy    188 atccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcaccgtggcg 247
          |   ||  ||| || | || |||| | || ||   || |||||||| |||| |   | | |
Db    109 aatgggtggtgatggtgacggcgcagactccaacgaatatcgccgtgattaagtattggg 168

Qy    248 aaaagcgcg 256
            |||| | |
Db    169 gaaagagag 177
```

RESULT 20
AAC33763
ID AAC33763 standard; DNA; 1541 BP.
XX
AC AAC33763;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 4228.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.8%; Score 37.8; DB 21; Length 1541;
Best Local Similarity 55.8%; Pred. No. 0.57;
Matches 72; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy    128 tgaaagtcaaggaaggcgatgccgtcaaaaaaggccaagtgctgtttgaagacaaaaaga 187
          | ||||  |   |   |  ||   || ||||||    ||   ||||    |   |  |||
Db     49 tcaaagctacacagatctctggaatcgaaaaagcttgagctgtgttaatggcggaggaga 108

Qy    188 atccgggcgtggtgtttactgcgccggcttcaggcaaaatcgccgcgattcaccgtggcg 247
          |   ||  ||| || | || |||| | || ||   || ||||||| |||| |   | | |
Db    109 aatgggtggtgatggtgacggcgcagactccaacgaatatcgccgtgattaagtattggg 168

Qy    248 aaaagcgcg 256
           |||| | |
Db    169 gaaagagag 177
```

RESULT 21
AAA81473/c
ID AAA81473 standard; DNA; 92934 BP.
XX
AC AAA81473;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_21 SEQ ID NO:21.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 471-498; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 92934 BP; 24051 A; 24249 C; 21902 G; 22730 T; 2 other;

Query Match 2.8%; Score 37.6; DB 21; Length 92934;
Best Local Similarity 52.6%; Pred. No. 4.8;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```
Qy    870 cgaattggttgacacagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
          ||||||||||  ||    | | |||    |  ||      | | | |||| ||||||
Db  64768 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 64709

Qy    930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaagg 989
          |   | |||   ||| | ||  ||   | || |  |||| ||      |  ||| |  |
Db  64708 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 64649

Qy    990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
          | ||| | || |    |    ||  | |||| | ||
Db  64648 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 64613
```

RESULT 22
AAF21613/c
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;

KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;

Query Match 2.8%; Score 37.6; DB 21; Length 172325;
Best Local Similarity 52.6%; Pred. No. 6.4;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 870 cgaattggttgacacagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
|||||||||| || | | ||| | || | | | |||| ||||||
Db 168435 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 168376

Qy 930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaagg 989
| | ||| ||| | || || | || | |||| || | ||| | |
Db 168375 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 168316

Qy 990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
| ||| | || | | || | |||| | ||
Db 168315 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 168280

RESULT 23

AAA81489/c
ID AAA81489 standard; DNA; 837096 BP.
XX
AC AAA81489;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 629-865; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Query Match 2.8%; Score 37.6; DB 21; Length 837096;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 870 cgaattggttgacacagacaaccgcgtgatttccggttcggtattgaacggcgcgattac 929
|||||||||| || | | ||| | || | | | |||| ||||||

```
Db 833206 CGAATTGGTTTCCATCCTCGAACGCCCTGTAACCAAAGAAGAAATCAACGCCGCGATGAA 833147

Qy    930 acaaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgtagaaagg 989
          |   | |||   ||| | ||  ||   | || |  |||| ||      |  ||| |  |
Db 833146 AGCCGCCGCCAGCGAATCTTACGGCTACAACGAAGATCAAATCGTTTCTTCCGACGTTGT 833087

Qy    990 ccgcagcaaagagctgttcggctgggttgcgccgca 1025
          | ||| | || |    |    ||  | |||| | ||
Db 833086 CGGCATCGAATACGGCTCGCTCTTCGATGCGACCCA 833051


RESULT  24
AAZ51001
ID   AAZ51001 standard; DNA; 1980 BP.
XX
AC   AAZ51001;
XX
DT   05-JUN-2000  (first entry)
XX
DE   X. oryzae avrRxo repeat domain DNA.
XX
KW   avrRxo gene; Xanthomonas oryzae pv. oryzicola; avirulence product;
KW   hypersensitive resistance; transgenic plant; cell death; abscission;
KW   fruit harvesting; herbicide development; ds.
XX
OS   Xanthomonas oryzae.
XX
FH   Key             Location/Qualifiers
FT   mat_peptide     1..1980
FT                   /*tag=  a
FT                   /product= "avrRxo repeat domain protein"
XX
PN   WO200009698-A2.
XX
PD   24-FEB-2000.
XX
PF   14-AUG-1999;   99WO-US18588.
XX
PR   14-AUG-1998;   98US-0096585.
PR   13-AUG-1999;   99US-0374273.
XX
PA   (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI   Leach JE,  Hulbert SH,  Ardales E;
XX
DR   WPI; 2000-224337/19.
DR   P-PSDB; AAY70063.
XX
PT   Novel avirulence gene from Xanthomonas oryzae, used to generate
PT   transgenic plants that show e.g. inducible resistance properties or
PT   cell death  -
XX
PS   Claim 1; Page 39-40; 46pp; English.
XX
CC   The present DNA sequence encodes avrRxo repeat domain from the rice
CC   pathogen, Xanthomonas oryzae pv. oryzicola. avrRxo gene encodes an
CC   avirulence product that can cause a hypersensitive resistance response
CC   in maize. Vectors containing avrRxo gene, preferably under the control
CC   of a tightly regulated inducible promoter, are used to produce
CC   transgenic plants that have inducible resistance properties. The
CC   hypersensivity response in plants like maize, that recognise the gene
CC   product, includes cell death, e.g. to prevent flowering, to induce
CC   abscission (to facilitate fruit harvesting), for herbicide development
CC   and for killing above-ground parts of plants before harvest.
XX
SQ   Sequence 1980 BP; 324 A; 622 C; 755 G; 279 T; 0 other;


  Query Match             2.8%;  Score 37;  DB 21;  Length 1980;
  Best Local Similarity   48.0%;  Pred. No. 1.1;
  Matches  106;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy   1008 cggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccattt 1067
          |||| |  |||| | || |||  |       |  ||       ||  | ||||
Db    960 cggcaagcaggcgctggagacggtgcagcggctgttgccggtgctgtgccaggcccatgg 1019

Qy   1068 cctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggt 1127
          |||||      | |   ||  |  || | |   |  |||||||||| | |  ||  |||
Db   1020 cctgaccccggaccaggtcgtggccatcgccaacaataacggcggcaagcaggcgctgga 1079

Qy   1128 gccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcg 1187
          | || |             ||  ||| ||  |  |||       |  ||||||     |
Db   1080 gacggtgcagcggctgttgccggtgctgtgccaggagcatggcctgaccctggaccaggt 1139

Qy   1188 cgatttaatcgtcggcgataccgacagcgcgcaggcattgg 1228
            |  |  |||| | || ||| || | ||  ||||||  |||
Db   1140 ggtgtccatcgccagcaataacggcggcaagcaggcgctgg 1180


RESULT  25
AAZ51000
ID   AAZ51000 standard; DNA; 3720 BP.
XX
AC   AAZ51000;
XX
DT   05-JUN-2000  (first entry)
XX
DE   X. oryzae avrRxo1 gene.
XX
KW   avrRxo1 gene; Xanthomonas oryzae pv. oryzicola; avirulence product;
KW   hypersensitive resistance; transgenic plant; cell death; abscission;
KW   fruit harvesting; herbicide development; ss.
XX
OS   Xanthomonas oryzae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3720
FT                   /*tag=  a
FT                   /product= "avrRxo protein"
XX
PN   WO200009698-A2.
XX
PD   24-FEB-2000.
XX
PF   14-AUG-1999;   99WO-US18588.
XX
PR   14-AUG-1998;   98US-0096585.
PR   13-AUG-1999;   99US-0374273.
XX
PA   (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI   Leach JE,  Hulbert SH,  Ardales E;
XX
DR   WPI; 2000-224337/19.
DR   P-PSDB; AAY70062.
XX
PT   Novel avirulence gene from Xanthomonas oryzae, used to generate
PT   transgenic plants that show e.g. inducible resistance properties or
PT   cell death  -
XX
PS   Claim 1; Page 36-37; 46pp; English.
XX
CC   The present sequence is avrRxo1 gene from the rice pathogen,
CC   Xanthomonas oryzae pv. oryzicola. This gene encodes an avirulence
CC   product that can cause a hypersensitive resistance response in maize.
CC   Vectors containing avrRxo gene, preferably under the control of a
CC   tightly regulated inducible promoter, are used to produce transgenic
CC   plants that have inducible resistance properties. The hypersensivity
CC   response in plants like maize, that recognise the gene product, includes
CC   cell death, e.g. to prevent flowering, to induce abscission (to
CC   facilitate fruit harvesting), for herbicide development and for killing
CC   above-ground parts of plants before harvest.
XX
SQ   Sequence 3720 BP; 646 A; 1192 C; 1295 G; 587 T; 0 other;
```

```
  Query Match             2.8%;  Score 37;  DB 21;  Length 3720;
  Best Local Similarity   48.0%;  Pred. No. 1.5;
  Matches  106;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy    1008 cggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccattt 1067
           ||||  |   |||| | || |||  |         |  ||        ||  | ||||
Db    1824 cggcaagcaggcgctggagacggtgcagcggctgttgccggtgctgtgccaggccatgg 1883

Qy    1068 cctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggt 1127
           |||||      | |   ||  |  || | |   |   |||||||| | |  ||  |||
Db    1884 cctgaccccggaccaggtcgtggccatcgccaacaataacggcggcaagcaggcgctgga 1943

Qy    1128 gccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcg 1187
           | || |          ||  ||| ||  |  |||       |  ||||||     |
Db    1944 gacggtgcagcggctgttgccggtgctgtgccaggagcatggcctgaccctggaccaggt 2003

Qy    1188 cgatttaatcgtcggcgataccgacagcgcgcaggcattgg 1228
           |  |  |||| | || ||| || | ||  ||||||  |||
Db    2004 ggtgtccatcgccagcaataacggcggcaagcaggcgctgg 2044
```

RESULT 26
AAZ95378
ID AAZ95378 standard; DNA; 1401 BP.
XX
AC AAZ95378;
XX
DT 01-JUN-2000 (first entry)
XX
DE Chlamydia pneumoniae antigen CPN100605 protein encoding DNA SEQ ID NO:1.
XX
KW Chlamydia pneumoniae; antigen; CPN100605 protein; immunisation;
KW vaccine; infection; antibacterial; antiinflammatory; bronchitis;
KW community acquired pneumonia; upper respiratory tract infection;
KW sinusitis; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 1..1401
FT /*tag= a
FT /product= "Antigen CPN100605 protein"
FT /note= "no stop codon given"
XX
PN WO200006742-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-IB01331.
XX
PR 27-JUL-1998; 98US-0094195.
PR 26-JUL-1999; 99US-0361443.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Murdin AD, Oomen RP;
XX
DR WPI; 2000-205466/18.
DR P-PSDB; AAY82082.
XX
PT Chlamydia pneumoniae antigens used for immunization and protection
PT against Chlamydia diseases -
XX
PS Claim 1; Fig 1; 48pp; English.
XX
CC The present sequence encodes the Chlamydia pneumoniae antigen CPN100605
CC protein. The CPN100605 protein has antibacterial and antiinflammatory
CC activities. The Chlamydia pneumoniae CPN100605 polynucleotide and protein
CC can be used in vaccination methods for preventing and treating Chlamydia
CC infection (e.g. infections caused by C. trachomatis, C. psittaci,
CC C. pneumoniae or C. pecorum). The polynucleotide can be used to produce
CC the protein recombinantly, in the construction of vaccine vectors,
CC as a vaccine agent, and in the construction of an attenuated Chlamydia
CC strain. The protein are also be useful as a vaccine agent, and for the
CC preparation of medicaments for treating or preventing Chlamydia
CC infection, e.g. community acquired pneumonia, and upper respiratory
CC tract infections such as bronchitis and sinusitis.
XX
SQ Sequence 1401 BP; 438 A; 308 C; 239 G; 416 T; 0 other;

```
  Query Match             2.7%;  Score 36.4;  DB 21;  Length 1401;
  Best Local Similarity   56.8%;  Pred. No. 1.5;
  Matches   67;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy     697 gcgaataaaaccgtgtggaccatcaattatcaagatgtaattaccattggccgtttgttt 756
           | |||  ||   || |  ||  |    | |||||||||  |||| |||||||  | | ||
Db     718 gagaaagaagtggtattcacactgtcatttcaagatgtccttactattggccatcttttc 777

Qy     757 gcaacaggccgtctgaacaccgagcgcgtgattgccctaggtggttctcaagtcaaca 814
             ||  ||  |  |      |||||  || |  || || | |||| ||  | | || |
Db     778 ttaaagggaagaattctgcacgagcaagtcacagcacttgctggtactgcactaaaaa 835
```

RESULT 27
AAC85030/c
ID AAC85030 standard; cDNA; 1725 BP.
XX
AC AAC85030;
XX
DT 08-MAY-2001 (first entry)
XX
DE Human Zcytor14 variant Zcytor14-1 degenerate nucleotide sequence.
XX
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine; variant; Zcytor14-1; ss.
XX
OS Homo sapiens.
XX
PN WO200104304-A1.
XX
PD 18-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18383.
XX
PR 07-JUL-1999; 99US-0348854.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Burkhead SK, Pownder SL;
XX
DR WPI; 2001-112618/12.
DR P-PSDB; AAB61881.
XX
PT New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis -
XX
PS Disclosure; Page 101; 112pp; English.
XX
CC The invention provides a new human cytokine receptor desinated Zcytor14.
CC Zcytor14 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
CC used to screen biological samples in vitro for the presence of Zcytor14.
CC Proteins, polypeptides and peptides having Zcytor14 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as rheumatoid
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
CC antibodies) can be used to treat a subject who produces an excess of
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
CC Zcytor14 to a subject. The present sequence represents a degenerate
CC nucleotide sequence of a variant of the human cytokine receptor Zcytor14,
CC designated Zcytor14-1. The variant is a truncated form of the receptor
CC polypeptide and lacks amino acid residues 1-113 of Zcytor14.
XX
SQ Sequence 1725 BP; 192 A; 226 C; 332 G; 216 T; 759 other;

Query Match 2.7%; Score 36.4; DB 22; Length 1725;
Best Local Similarity 25.0%; Pred. No. 1.6;
Matches 82; Conservative 47; Mismatches 199; Indels 0; Gaps 0;

```
Qy   1008 cggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccattt 1067
          |: | ||:|   :  :  :   :| : |  | ||  |:   :: |      |     :|
Db   1489 CRTCNGGRTGNARNARNCKRTCRAARCANGCNCCNACRTANSWNCCNGGNGCNCKNCCYT 1430

Qy   1068 cctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggt 1127
             ::||: |        |:      |  ::  |   :|  ||: |:   |  ||:||
Db   1429 GNARRAARTCNGGNARNACRCANSWNARNSWNGCNCKRAANGCRTCRTGNGGNCCRTGNG 1370

Qy   1128 gccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcg 1187
             |       :: ||     | ||  :|| :|   ::    :  |  |:           | :
Db   1369 CNCCNGGNCCNSWNACNCCRTCYTGNARCCAYTCNSWRCANARNGCNACNGCNCCNGGNS 1310

Qy   1188 cgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacga 1247
          :|   :     | | |  ||  |  |  | |:   :||  :   :||    .||:|
Db   1309 WRAANARNARNACNACNACNCCNCCYTCYTGNARNGTYTGNCKNCKYTGNGCRTGRAACC 1250

Qy   1248 agaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgtt 1307
          |    ||    | :|| || :     ||    |: :   | | :   || ||  :
Db   1249 ANGCNACNGGNCCYTGNGCNSWNARYTCNCKNCKNSWCCANARRTCNACNGCNACNCKNA 1190

Qy   1308 gcgcaaagtgctggaaaccattgagaag 1335
          : |  |: ||    |:  |  :   |:
Db   1189 RNGGNARYTGRCANARNGCNSWNGCNAR 1162
```

RESULT 28
AAC85028/c
ID AAC85028 standard; cDNA; 2076 BP.
XX
AC AAC85028;
XX
DT 08-MAY-2001 (first entry)
XX
DE Human cytokine receptor Zcytor14 degenerate nucleotide sequence.
XX
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;
KW antiinflammatory; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200104304-A1.
XX
PD 18-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18383.
XX
PR 07-JUL-1999; 99US-0348854.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Burkhead SK, Pownder SL;
XX
DR WPI; 2001-112618/12.
DR P-PSDB; AAB61880.
XX
PT New polypeptide encoding a human cytokine receptor Zcytor14, for
PT treating inflammation e.g. rheumatoid arthritis -
XX
PS Claim 8; Page 94-95; 112pp; English.
XX
CC The invention provides a new human cytokine receptor desinated Zcytor14.
CC Zcytor14 can be expressed by standard recombinant methodology. The
CC encoding nucleic acid is useful for detecting the expression of a
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be
CC used to screen biological samples in vitro for the presence of Zcytor14.
CC Proteins, polypeptides and peptides having Zcytor14 activity can be
CC administered to a subject who lacks an adequate amount of this
CC polypeptide, for treating inflammation and conditions such as rheumatoid
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14
CC antibodies) can be used to treat a subject who produces an excess of
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide
CC Zcytor14 to a subject. The present sequence represents the human
CC cytokine receptor Zcytor14 degenerate nucleotide sequence.
XX
SQ Sequence 2076 BP; 232 A; 273 C; 399 G; 265 T; 907 other;

Query Match 2.7%; Score 36.4; DB 22; Length 2076;
Best Local Similarity 25.0%; Pred. No. 1.8;
Matches 82; Conservative 47; Mismatches 199; Indels 0; Gaps 0;

```
Qy   1008 cggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccattt 1067
          |: | ||:|   :  :  :   :| : |  | ||  |:   :: |      |     :|
Db   1840 CRTCNGGRTGNARNARNCKRTCRAARCANGCNCCNACRTANSWNCCNGGNGCNCKNCCYT 1781

Qy   1068 cctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggt 1127
             ::||: |        |:      |  ::  |   :|  ||: |:   |  ||:||
Db   1780 GNARRAARTCNGGNARNACRCANSWNARNSWNGCNCKRAANGCRTCRTGNGGNCCRTGNG 1721

Qy   1128 gccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcg 1187
             |       :: ||     | ||  :|| :|   ::    :  |  |:           | :
Db   1720 CNCCNGGNCCNSWNACNCCRTCYTGNARCCAYTCNSWRCANARNGCNACNGCNCCNGGNS 1661

Qy   1188 cgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacga 1247
             :|   :     | | |  ||  |  |  | |:   :||  :   :||    ||:|
Db   1660 WRAANARNARNACNACNACNCCNCCYTCYTGNARNGTYTGNCKNCKYTGNGCRTGRAACC 1601

Qy   1248 agaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgtt 1307
          |    ||    | :|| || :     ||    |: :   | | :   || ||     :
Db   1600 ANGCNACNGGNCCYTGNGCNSWNARYTCNCKNCKNSWCCANARRTCNACNGCNACNCKNA 1541

Qy   1308 gcgcaaagtgctggaaaccattgagaag 1335
          : |  |: ||    |:  |  :   |:
Db   1540 RNGGNARYTGRCANARNGCNSWNGCNAR 1513
```

RESULT 29
AAS18131/c
ID AAS18131 standard; cDNA; 2109 BP.
XX
AC AAS18131;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human DCRS7 reverse translation generic cDNA.
XX
KW Human; DNAX cytokine receptor subunit 7; DCRS7; phosphate labelling; ss;
KW gene therapy; protein therapy; immunological disorder.
XX
OS Homo sapiens.
XX
PN WO200190358-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US16767.
XX
PR 24-MAY-2000; 2000US-206862P.
XX
PA (SCHE) SCHERING CORP.
XX
PI Gorman DM;
XX
DR WPI; 2002-106198/14.
XX
PT Isolated antigenic human or mouse DNAX receptor subunit-like
PT polypeptide useful for detecting antibodies generated in response to
PT presence of increased protein levels or immunological disorders -
XX
PS Disclosure; Page 15-16; 148pp; English.

```
XX
CC   The invention relates to primate and rodent DNAX cytokine receptor
CC   subunit (DCRS) polypeptides and the polynucleotides encoding them. The
CC   receptors, or their portions may be useful as phosphate labelling enzymes
CC   to label general or specific substrates. The subunits may also be
CC   functional immunogens to elicit recognising antibodies, or antigens
CC   capable of binding antibodies. A combination, e.g., including a DCRS can
CC   be used as an immunogen for the production of antisera or antibodies
CC   capable of distinguishing between other cytokine receptor family members.
CC   A purified DCRS can also be used as a reagent to detect antibodies
CC   generated in response to the presence of elevated levels of expression,
CC   or immunological disorders which lead to antibody production to the
CC   endogenous receptor. This sequence represents human DCRS7 reverse
CC   translation generic cDNA.
XX
SQ   Sequence 2109 BP; 234 A; 267 C; 415 G; 269 T; 924 other;

  Query Match             2.7%;  Score 36.4;  DB 24;  Length 2109;
  Best Local Similarity   25.0%;  Pred. No. 1.8;
  Matches   82;  Conservative   47;  Mismatches  199;  Indels    0;  Gaps    0;

Qy    1008 cggctgggttgcgccgcagccggacaaatactccatcacgcgtacaaccctcggccattt 1067
           |: | ||:|   :  :  :   :| : |  | ||  |:   :: |     |       :|
Db    1873 CRTCNGGRTGNARNARNCKRTCRAARCANGCNCCNACRTANSWNCCNGGNGCNCKNCCYT 1814

Qy    1068 cctgaaaaacaaactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggt 1127
             ::||: |        |:      |  ::  |   :|  ||: |:    |  ||:||
Db    1813 GNARRAARTCNGGNARNACRCANSWNARNSWNGCNCKRAANGCRTCRTGNGGNCCRTGNG 1754

Qy    1128 gccgattggtacttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcg 1187
             |         :: ||      | ||  :|| :|   ::     :  |  |         | :
Db    1753 CNCCNGGNCCNSWNACNCCRTCYTGNARCCAYTCNSWRCANARNGCNACNGCNCCNGGNS 1694

Qy    1188 cgatttaatcgtcggcgataccgacagcgcgcaggcattgggttgcttggaattggacga 1247
            :|   :     |  | |  ||  |  |  | |:   :||  :   :||.    ||:|
Db    1693 WRAANARNARNACNACNACNCCNCCYTCYTGNARNGTYTGNCKNCKYTGNGCRTGRAACC 1634

Qy    1248 agaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctgtt 1307
           |   ||    | :|| || :     ||    |: :   | | :   || ||     :
Db    1633 ANGCNACNGGNCCYTGNGCNSWNARYTCNCKNCKNSWCCANARRTCNACNGCNACNCKNA 1574

Qy    1308 gcgcaaagtgctggaaaccattgagaag 1335
           : |  |: ||    |:  |  :    |:
Db    1573 RNGGNARYTGRCANARNGCNSWNGCNAR 1546


RESULT  30
AAD17186
ID   AAD17186 standard; DNA; 125401 BP.
XX
AC   AAD17186;
XX
DT   29-NOV-2001  (first entry)
XX
DE   Streptomyces noursei nystatin PKS gene cluster DNA.
XX
KW   Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW   antifungal; antibiotic; ds.
XX
OS   Streptomyces noursei.
XX
FH   Key             Location/Qualifiers
FT   CDS             6337..34771
FT                   /*tag=  a
FT                   /product= "NysI complete protein"
FT   CDS             34792..51099
FT                   /*tag=  b
FT                   /product= "NysJ protein"
FT   CDS             51155..57355
FT                   /*tag=  c
FT                   /product= "NysK protein"
FT   CDS             57503..58687
FT                   /*tag=  d
FT                   /product= "NysL protein"
FT   CDS             complement (58786..58980)
FT                   /*tag=  e
FT                   /product= "NysM protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (59045..60241)
FT                   /*tag=  f
FT                   /product= "NysN protein"
FT                   /note= "CDS does not include start codon"
FT   CDS             complement (60238..61296)
FT                   /*tag=  g
FT                   /product= "NysD2 complete protein"
FT   CDS             120628..121308
FT                   /*tag=  h
FT                   /product= "NysR4 (long) protein"
XX
PN   WO200159126-A2.
XX
PD   16-AUG-2001.
XX
PF   08-FEB-2001; 2001WO-GB00509.
XX
PR   08-FEB-2000; 2000GB-0002840.
PR   10-APR-2000; 2000GB-0008786.
PR   14-APR-2000; 2000GB-0009387.
XX
PA   (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA   (SNTF ) SINTEF STIFTELSEN IND TEK FORSK.
PA   (ALPH-) ALPHARMA AS.
PA   (SINV-) SINVENT AS.
PA   (DZIE/) DZIEGLEWSKA H.
PA   (ZOTC/) ZOTCHEV S B.
PA   (SEKU/) SEKUROVA O N.
PA   (FJAE/) FJAERVIK E.
PA   (BRAU/) BRAUTASET T.
PA   (STRO/) STROM A R.
XX
PI   Zotchev SB,  Sekurova ON,  Fjaervik E,  Brautaset T,  Strom AR;
PI   Valla S,  Ellingsen TE,  Sletta H,  Gulliksen O;
XX
DR   WPI; 2001-557614/62.
DR   P-PSDB; AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
DR   AAE10149, AAE10150.
XX
PT   New nystatin polyketide synthase polynucleotides and polypeptides,
PT   useful as antibiotics and antifungals -
XX
PS   Claim 1; Page 188-254; 266pp; English.
XX
CC   The present invention relates to the cloning and sequencing of the gene
CC   cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC   involved in the biosynthesis of the macrolide antibiotic nystatin.
CC   The nystatin PKS is useful as antifungal antibiotics. The present
CC   sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
XX
SQ   Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;

  Query Match             2.6%;  Score 35.6;  DB 22;  Length 125401;
  Best Local Similarity   47.7%;  Pred. No. 23;
  Matches  104;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy    1104 caacggcggcgaccgcgccatggtgccgattggtacttacgagcgcgtgatgcccttgga 1163
           || || ||||| || || |   |   |  | || ||      || || |  ||
Db   50040 caccgccggcgtcctcgacgacgccaccgtcggcaccctgacgcccgagcggctggacac 50099

Qy    1164 tatcctgcccaccctgcttttgcgcgatttaatcgtcggcgataccgacagcgcgcaggc 1223
             |||||| | ||  |        ||  |    | |   |||  ||  | |||  | |
Db   50100 cgtcctgcgcgccaaggcggacgccgcctggcacctgcacgacgccacccgcgaccgcga 50159

Qy    1224 attgggttgcttggaattggacgaagaagacctcgctttgtgcagcttcgtctgcccggg 1283
             |||   | || |   || ||      | |  || | |   | ||  |  | ||| |||
```

```
Db  50160 cctggcagggttcgtgctgtactcctcggtcgccggtgtcaccggcggcccggccaggg 50219

Qy   1284 caaatacgaatacggcccgctgttgcgcaaagtgctgg 1321
          ||| ||||    ||||     ||| | | | | ||| |
Db  50220 caactacgccgccggcaacacgttcctcgacgcgctcg 50257
```

RESULT 31
AAS57519
ID AAS57519 standard; cDNA; 296 BP.
XX
AC AAS57519;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #195 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US09246.
XX
PR 24-MAR-2000; 2000US-191597P.
PR 04-MAY-2000; 2000US-202024P.
PR 05-MAY-2000; 2000US-202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
DR WPI; 2001-611627/70.
XX
PT New colon tumour proteins and related nucleic acid, useful for
PT treatment, prevention, diagnosis and monitoring of cancer -
XX
PS Claim 4; Page 90; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA sequences
CC encoding for at least an immunogenic portion of human colon tumour
CC proteins. The sequences of the invention are useful in pharmaceutical
CC compositions and vaccines for the prevention and treatment of cancers
CC such as colon cancer. They are also useful for the diagnosis and
CC monitoring of such cancers. Antibodies to the colon tumour proteins
CC and antigen presenting cells that express polynucleotides encoding
CC colon tumour proteins can be used to inhibit the development of
CC cancers. T-cells that react specifically with colon tumour proteins
CC are useful for removing tumour cells from samples (e.g. blood) and
CC for cancer treatment. The polynucleotides sequences are also useful in
CC gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
CC invention that encode for portions of human colon tumour proteins.
XX
SQ Sequence 296 BP; 66 A; 86 C; 88 G; 51 T; 5 other;

```
  Query Match              2.6%;  Score 35.4;  DB 23;  Length 296;
  Best Local Similarity   54.8%;  Pred. No. 1.4;
  Matches   69;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy   1215 cgcgcaggcattgggttgcttggaattggacgaagaagacctcgctttgtgcagcttcgt 1274
           || || | |    ||     |   |  || ||  || ||    || |   || || |
Db     50 cgagcggccgcccgggcaggtctcangggccgccgaggatggggcctgcggctgcctgcc 109

Qy   1275 ctgcccgggcaaatacgaatacggcccgctgttgcgcaaagtgctggaaaccattgagaa 1334
          |  |||||||| || |||  |  |||  | |  | |||| | ||||    ||||| |||
Db    110 caacccgggcacattcgaggagtgccaccggaagtgcaaggagctgtttcccattcagat 169

Qy   1335 ggaagg 1340
          ||| ||
Db    170 ggaggg 175
```

RESULT 32
AAZ09101/c
ID AAZ09101 standard; DNA; 763 BP.
XX
AC AAZ09101;
XX
DT 18-OCT-1999 (first entry)
XX
DE Mistletoe lectin A-chain MLA DNA consensus sequence.
XX
KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW cancer; cytotoxicity; antigen; isoform; ss.
XX
OS Viscum album.
XX
PN DE19804210-A1.
XX
PD 12-AUG-1999.
XX
PF 03-FEB-1998; 98DE-1004210.
XX
PR 03-FEB-1998; 98DE-1004210.
XX
PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI Morris P, Stiefel T, Voelter W, Welters P;
XX
DR WPI; 1999-445335/38.
XX
PT Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
XX
PS Claim 12; Page 34; 78pp; German.
XX
CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
CC described in the specification.
XX
SQ Sequence 763 BP; 163 A; 215 C; 175 G; 170 T; 40 other;

```
  Query Match              2.6%;  Score 35.4;  DB 20;  Length 763;
  Best Local Similarity   47.0%;  Pred. No. 2.2;
  Matches   78;  Conservative    8;  Mismatches   80;  Indels    0;  Gaps    0;

Qy    871 gaattggttgacacagacaaccgcgtgatttccggttcggtattgaacggcgcgattaca 930
          || | || ||   | |   |    |||  || | |  | |  | |  || ||||   | |
Db    321 GANTCGGGTGGTGCCGGTGAAGAGGTGCGTTTCCGCGCCGCGTGGTGCGTCGCGCAAAAA 262

Qy    931 caaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaaggc 990
            |||   |  ||  |  |||| | || || ||   :::|||       |||| |  |||
Db    261 GTAGGATTGGTCGCCTGCTTGGTAAGCCACGACGTAMRSATTGGTAACGTCGATGGCGGC 202

Qy    991 cgcagcaaagagctgttcggctgggttgcgccgcagccggacaaat 1036
          || : :  ||  :::  |      | || ||  || |  |||||||
Db    201 CGTRAYCGAGTCYYSCCCCTGGTTGGTGAGCTCCACCAAGACAAAT 156
```

```
RESULT  33
AAZ09100/c
ID   AAZ09100 standard; DNA; 1598 BP.
XX
AC   AAZ09100;
XX
DT   18-OCT-1999  (first entry)
XX
DE   Mistletoe lectin DNA consensus sequence.
XX
KW   Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW   ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW   lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW   cancer; cytotoxicity; antigen; isoform; ss.
XX
OS   Viscum album.
XX
PN   DE19804210-A1.
XX
PD   12-AUG-1999.
XX
PF   03-FEB-1998;   98DE-1004210.
XX
PR   03-FEB-1998;   98DE-1004210.
XX
PA   (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI   Morris P,  Stiefel T,  Voelter W,  Welters P;
XX
DR   WPI; 1999-445335/38.
XX
PT   Preparation of mistletoe lectins in heterologous systems,
PT   particularly for use as anticancer agents and immunostimulants
XX
PS   Claim 11; Page 32-33; 78pp; German.
XX
CC   This invention describes a novel mistletoe lectin (I) and its fragments
CC   which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC   of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC   ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC   lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC   fragments are used to treat uncontrolled cell growth (particularly
CC   cancers) and if they lack cytotoxicity, to increase the strength of the
CC   immune response, particularly to a co-administered antigen
CC   (tumour-associated, bacterial or viral). The method allows production of
CC   mistletoe lectin, and its individual chains, in many different isoforms
CC   and on a large scale, at any time of the year. Recombinant products are
CC   free from toxins present in natural mistletoe extracts. This sequence
CC   represents a consensus sequence of mistletoe lectin DNA described in the
CC   specification.
XX
SQ   Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;


  Query Match             2.6%;  Score 35.4;  DB 20;  Length 1598;
  Best Local Similarity   47.0%;  Pred. No. 3.2;
  Matches   78;  Conservative    8;  Mismatches   80;  Indels    0;  Gaps    0;

Qy    871 gaattggttgacacagacaaccgcgtgatttccggttcggtattgaacggcgcgattaca 930
          || | || ||   | |   |    |||  || | |  | |  | |  || ||||   | |
Db    321 GANTCGGGTGGTGCCGGTGAAGAGGTGCGTTTCCGCGCCGCGTGGTGCGTCGCGCAAAAA 262

Qy    931 caaggcgcgcacgattatttgggacgctaccacaatcagatttccgttatcgaagaaggc 990
            |||   |  ||  |  |||| | || || ||   :::|||       |||| |  |||
Db    261 GTAGGATTGGTCGCCTGCTTGGTAAGCCACGACGTAMRSATTGGTAACGTCGATGGCGGC 202

Qy    991 cgcagcaaagagctgttcggctgggttgcgccgcagccggacaaat 1036
          || : :  ||  :::  |      | || ||  || |  |||||||
Db    201 CGTRAYCGAGTCYYSCCCCTGGTTGGTGAGCTCCACCAAGACAAAT 156


RESULT  34
```

```
AAZ32025
ID   AAZ32025 standard; DNA; 9810 BP.
XX
AC   AAZ32025;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH1 related EST AF018073.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 353-359; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 9810 BP; 1583 A; 3401 C; 3201 G; 1625 T; 0 other;


  Query Match             2.6%;  Score 35.4;  DB 20;  Length 9810;
  Best Local Similarity   50.9%;  Pred. No. 7.6;
  Matches   84;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy     990 ccgcagcaaagagctgttcggctgggttgcgccgcagccggacaaatactccatcacgcg 1049
           |||||   | | |||  |   |  | |     ||||||    | |   | || ||  | |
Db    1852 ccgcatgcaggggctcatgtccgagttcaacgcgcagcaccccgacatcaccgtcgagtg 1911

Qy    1050 tacaaccctcggccatttcctgaaaaacaaactcttcaagttcaacacagccgtcaacgg 1109
             || ||||   |   | ||     | |     ||| |  | |||  |||  ||| ||
```

```
Db   1912 ggtgacgctcgaggaaaacgtgctgcgccagaaggtcacgaccgacatcgccaccaaggg 1971

Qy   1110 cggcgaccgcgccatggtgccgattggtacttacgagcgcgtgat 1154
          |||  |   || | || || | || || || |||||     |||
Db   1972 cgggcagttcgacgtgctgaccatcggcacctacgaggttccgat 2016


RESULT 35
AAC90082
ID   AAC90082 standard; DNA; 9810 BP.
XX
AC   AAC90082;
XX
DT   19-MAR-2001  (first entry)
XX
DE   AF018073 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 7; Pages 653-659; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 9810 BP; 1583 A; 3401 C; 3201 G; 1625 T; 0 other;


  Query Match             2.6%;  Score 35.4;  DB 22;  Length 9810;
  Best Local Similarity   50.9%;  Pred. No. 7.6;
  Matches   84;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy    990 ccgcagcaaagagctgttcggctgggttgcgccgcagccggacaaatactccatcacgcg 1049
          |||||   | | |||  |   |  | |     ||||||    | |   | || ||  | |
Db   1852 ccgcatgcaggggctcatgtccgagttcaacgcgcagcaccccgacatcaccgtcgagtg 1911

Qy   1050 tacaaccctcggccatttcctgaaaaacaaactcttcaagttcaacacagccgtcaacgg 1109
            || ||||   |   | ||     | |     ||| |  | |||  |||  ||| ||
Db   1912 ggtgacgctcgaggaaaacgtgctgcgccagaaggtcacgaccgacatcgccaccaaggg 1971

Qy   1110 cggcgaccgcgccatggtgccgattggtacttacgagcgcgtgat 1154
          |||  |   || | || || | || || || |||||     |||
Db   1972 cgggcagttcgacgtgctgaccatcggcacctacgaggttccgat 2016


RESULT 36
AAS86095
ID   AAS86095 standard; cDNA; 323 BP.
XX
AC   AAS86095;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21899.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21908.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 21899; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
```

```
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 323 BP; 95 A; 86 C; 84 G; 58 T; 0 other;


  Query Match             2.6%;  Score 35;  DB 23;  Length 323;
  Best Local Similarity   59.6%;  Pred. No. 1.9;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
Db     37 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 96

Qy   1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
             | |  | |||| | ||||    ||||| ||| ||| ||
Db     97 ccggaagtgcaaggagctgtttcccattcagatggaggg 135


RESULT  37
AAF21735
ID   AAF21735 standard; DNA; 1139 BP.
XX
AC   AAF21735;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 122.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; AAB58832.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 573; 1299pp; English.
XX
CC   Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC   proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC   associated with breast and ovarian cancer. Included in the invention are
CC   sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC   isolation and characterisation of the DNA and protein sequences of the
CC   invention. The breast and ovarian cancer associated DNA, protein, agonist
CC   or antagonist sequences exhibit cytostatic; immunosuppressive;
CC   nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC   antibacterial; antifungal; antiparasitic and cardiant activity. The
CC   polynucleotide and protein sequences are used in the diagnosis of cancer,
CC   particularly breast and ovarian cancer. The nucleic acid sequences,
CC   proteins, agonists and agonists may also be used in the diagnosis,
CC   prevention and treatment of immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ   Sequence 1139 BP; 217 A; 350 C; 375 G; 188 T; 9 other;


  Query Match             2.6%;  Score 35;  DB 21;  Length 1139;
  Best Local Similarity   59.6%;  Pred. No. 3.6;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
           || ||  || ||    || |   || || |   |  |||||||| || |||  |  |||
Db    353 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 412

Qy   1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
             | |  | |||| | ||||    ||||| ||| ||| ||
Db    413 ccggaagtgcaaggagctgtttcccattcagatggaggg 451


RESULT  38
AAS86098
ID   AAS86098 standard; cDNA; 1689 BP.
XX
AC   AAS86098;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #21902.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG21911.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 21902; 103pp; English.
```

```
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1689 BP; 348 A; 538 C; 525 G; 278 T; 0 other;


  Query Match             2.6%;  Score 35;  DB 23;  Length 1689;
  Best Local Similarity   59.6%;  Pred. No. 4.3;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy   1242 ggacgaagaagacctcgctttgtgcagcttcgtctgcccgggcaaatacgaatacggccc 1301
          || ||  || ||    || |   || || |   |  |||||||| || |||  | |||
Db    285 ggccgccgaggatggggcctgcggctgcctgcccaacccgggcacattcgaggagtgcca 344

Qy   1302 gctgttgcgcaaagtgctggaaaccattgagaaggaagg 1340
          | |  | |||| | ||||    ||||| ||| ||| ||
Db    345 ccggaagtgcaaggagctgtttcccattcagatggaggg 383


RESULT  39
AAN60550
ID   AAN60550 standard; cDNA; 4383 BP.
XX
AC   AAN60550;
XX
DT   01-JAN-1980  (first entry)
XX
DE   Mouse erythroid Band 3 anion transport protein.
XX
KW   Band 3; anion transport protein; glycoprotein; cystic fibrosis;
KW   disease diagnosis; erythrocyte; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             127..2915
FT                   /*tag=  a
XX
PN   WO8605203-A.
XX
PD   12-SEP-1986.
XX
PF   06-MAR-1986;   86WO-US00450.
XX
PR   07-MAR-1985;   85US-0709183.
XX
PA   (WHIT-) WHITEHEAD INST BIOM.
XX
PI   Alper SI,  Kopito RR;
XX
DR   WPI; 1986-252231/38.
DR   P-PSDB; AAP60644.
XX
PT   DNA coding for anion transport protein - esp. for detection of
PT   genetic defects
XX
PS   Disclosure; Fig. 2; 49pp; English.
XX
CC   Band 3 is the major glycoprotein of mammalian erythrocytes and
CC   functions as an anion antiport in that it mediates the 1-for-1
CC   exchange of chloride (Cl-) and bicarbonate (HCO3-) across the
CC   erythrocyte plasma membrane.  The anion transport protein may be
CC   used in the detection of genetic defects resulting in abnormal
CC   transport of ions across cell membranes, e.g. as in cystic fibrosis.
CC   This cDNA is specifically from mouse erythrocytes.  See also AAN60551
CC   and AAP60645.
XX
SQ   Sequence 4383 BP; 960 A; 1257 C; 1136 G; 1030 T; 0 other;


  Query Match             2.6%;  Score 35;  DB 7;  Length 4383;
  Best Local Similarity   47.1%;  Pred. No. 6.8;
  Matches  107;  Conservative    0;  Mismatches  120;  Indels    0;  Gaps    0;

Qy   1079 aactcttcaagttcaacacagccgtcaacggcggcgaccgcgccatggtgccgattggta 1138
          | ||||||| |  || |     | ||  |   |  |  |      |||||  |     |
Db   2681 acctcttcacgggcatccagatcatctgcctggctgtgctgtgggtggtgaagtccactc 2740

Qy   1139 cttacgagcgcgtgatgcccttggatatcctgcccaccctgcttttgcgcgatttaatcg 1198
          ||  |  ||  |   |||||||  |   || | | |||  ||| | | ||   | | |||
Db   2741 ctgcctcgctggccttgcccttcgtcctcatcctcacagtgcctctccgtcgtctcatcc 2800

Qy   1199 tcggcgataccgacagcgcgcaggcattgggttgcttggaattggacgaagaagacctcg 1258
          ||      | |  ||| | || || | |    ||  ||||    || || |   |  |
Db   2801 tcccgctcatcttcagagagctggaactccagtgtctggacggtgatgatgccaaagtga 2860

Qy   1259 ctttgtgcagcttcgtctgcccgggcaaatacgaatacggcccgctg 1305
          | ||   |          ||| ||   |||| ||  | |  ||  ||
Db   2861 cctttgacgaggagaatggcctggatgaatatgacgaagtgcccatg 2907


RESULT  40
AAD08837
ID   AAD08837 standard; cDNA; 1325 BP.
XX
AC   AAD08837;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Human G-protein coupled receptor-3 (GCREC-3) cDNA.
XX
KW   Human; G-protein coupled receptor-3; GCREC-3; gene therapy; cirrhosis;
KW   transgenic animal; proliferative disorder; actinic keratosis; hepatitis
KW   nephrotropic; cancer; breast; bladder; bone marrow; brain; uterus;
KW   leukaemia; adenocarcinoma; lymphoma; melanoma; myeloma; epilepsy; stroke;
KW   neurological disorder; Alzheimer's disease; Parkinson's disease; nausea;
KW   Huntington's disease; multiple sclerosis; dementia; angina pectoris;
KW   central nervous system disorder; cardiovascular disorder; hypertension;
KW   atherosclerosis; congestive heart failure; gastrointestinal disorder;
KW   dysphagia; peptic oesophagitis; spasm; gastritis; anorexia; pyrosis;
KW   pancreatitis; Crohn's disease; diarrhoea; autoimmune disorder; anaemia;
KW   inflammatory disorder; Acquired Immune Deficiency Syndrome; AIDS;
KW   Addison's diseases; allergy; asthma; diabetes mellitus; antithyroid;
KW   atopic dermatitis; glomerulonephritis; Grave's disease; psoriasis;
KW   rheumatoid arthritis; ulcerative colitis; osteoporosis; antifungal;
KW   metabolic disorder; obesity; nootropic; protozoacide; virucide; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             42..1310
FT                   /*tag=  a
FT                   /product= "Human GCREC-3 protein"
XX
PN   WO200142288-A2.
XX
```

PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33382.
XX
PR 10-DEC-1999; 99US-0172852.
PR 22-DEC-1999; 99US-0171732.
PR 14-JAN-2000; 2000US-0176148.
PR 21-JAN-2000; 2000US-0177331.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Burford N, Baughn MR, Au-Young J, Yang J, Lu DAM, Reddy R;
XX
DR WPI; 2001-381635/40.
DR P-PSDB; AAE04547.
XX
PT New human G-protein coupled receptor polypeptides for diagnosing,
PT preventing, and treating cell proliferative, neurological,
PT cardiovascular, gastrointestinal, autoimmune and metabolic disorders -
XX
PS Claim 5; Page 160-161; 175pp; English.
XX
CC The present sequence is human G-protein coupled receptor-3 (GCREC-3)
CC cDNA. GCREC is useful in somatic or germline gene therapy to correct a
CC genetic deficiency, to express a conditionally lethal gene product and
CC to express a protein which affords protection against intracellular
CC parasites and also for diagnosis of disorders associated with expression
CC of GCREC. GCREC is also useful for generating hybridisation probes useful
CC in mapping the naturally occurring genomic sequences and to create
CC knockin humanised animals (pigs) or transgenic animals (mice or rats) to
CC model human diseases. GCREC is used to diagnose, prevent and treat
CC proliferative disorders (actinic keratosis, arteriosclerosis, cirrhosis,
CC hepatitis and cancer); cancer (breast, bladder, bone marrow, brain,
CC uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma)
CC neurological disorders (epilepsy, stroke, Alzheimer's, Huntington's,
CC Parkinson's disease, multiple sclerosis, dementia and other central
CC nervous system disorders); cardiovascular disorders (angina pectoris,
CC hypertension, atherosclerosis, congestive heart failure);
CC gastrointestinal disorders (dysphagia, peptic oesophagitis, oesophageal
CC spasm, gastritis, gastric carcinoma, anorexia, nausea, abdominal angina,
CC pyrosis, pancreatitis, Crohn's disease, diarrhoea); autoimmune/
CC inflammatory disorders (acquired immunodeficiency syndrome (AIDS),
CC Addison's diseases, allergies, anaemia, asthma, diabetes mellitus, atopic
CC dermatitis, glomerulonephritis, Grave's disease, osteoarthritis,
CC psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal,
CC parasitic, protozoal and helminthic infections) and metabolic disorders
CC (obesity, osteoporosis, viral infections).
XX
SQ Sequence 1325 BP; 195 A; 473 C; 403 G; 254 T; 0 other;

Query Match 2.6%; Score 34.8; DB 22; Length 1325;
Best Local Similarity 51.3%; Pred. No. 4.4;
Matches 81; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 973 tccgttatcgaagaaggccgcagcaaagagctgttcggctgggttgcgccgcagccggac 1032
Db 933 tactttagcgccgagggcttccgcaacaccctgcgcggcctgggcactccgcaccgggcc 992

Qy 1033 aaatactccatcacgcgtacaaccctcggccatttcctgaaaaacaaactcttcaagttc 1092
Db 993 aggacctcggccaccaacgggacgcgggcggcgctcgcgcaatccgaaaggtccgccgtc 1052

Qy 1093 aacacagccgtcaacggcggcgaccgcgccatggtgcc 1130
Db 1053 accaccgacgccaccaggccggatgccgccatgtcccc 1090

RESULT 41
AAF25080/c
ID AAF25080 standard; DNA; 697 BP.
XX
AC AAF25080;
XX
DT 30-APR-2001 (first entry)
XX
DE Nucleotide sequence of a M. vaccae immunogenic epitope from DNA9A.
XX
KW Epitope; antigen; cytokine production; immune disorder; tuberculosis;
KW cancer; mycobacterial infection; TH1 immune response; vaccine; ss.
XX
OS Mycobacterium vaccae.
XX
FH Key Location/Qualifiers
FT CDS 1..225
FT /*tag= a
FT /note= "no termination sequence given"
XX
PN WO200104140-A1.
XX
PD 18-JAN-2001.
XX
PF 10-JUL-2000; 2000WO-NZ00121.
XX
PR 12-JUL-1999; 99US-0351348.
PR 29-NOV-1999; 99US-0450072.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Delcayre A;
XX
DR WPI; 2001-168411/17.
DR P-PSDB; AAB31649.
XX
PT Novel polypeptides comprising immunogenic epitopes of Mycobacterium
PT vaccae, useful for treating mycobacterial infections, immune disorders
PT and cancers -
XX
PS Claim 4; Page 59; 80pp; English.
XX
CC The specification describes an immunogenic epitope of a Mycobacterium
CC vaccae antigen. The epitope is a stimulator of cytokine production.
CC The epitopes are useful for the treatment of immune disorders,
CC infectious diseases, especially tuberculosis, and cancer. They are
CC also useful for treatment of other mycobacterial infections such as
CC those caused by Mycobacterium avium. The epitopes are especially useful
CC for inducing TH1 immune responses, and for producing vaccines. The
CC present sequence encodes a M. vaccae epitope of the invention.
XX
SQ Sequence 697 BP; 89 A; 253 C; 253 G; 102 T; 0 other;

Query Match 2.6%; Score 34.6; DB 22; Length 697;
Best Local Similarity 46.5%; Pred. No. 3.7;
Matches 112; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 277 gccgttgaaggcaacgacgaaatcgagtttgaacgctacgcacctgaagcgctggcaaac 336
Db 447 GCCGTGCGCGGCGCCGACGCCGACGCGTACGCCGCCGCGGTTCGGGATGCGCTGGCACTG 388

Qy 337 ttaagcggcgaagaagtgcgccgcaacctgatccaatccggtttgtggactgcgctgcgc 396
Db 387 ATCGAAGAGGCCGGACAGCGGGCGGTCCGCCTGCCCGGCGACCTGGCCGAGCCGGCGCAC 328

Qy 397 acccgtccgttcagcaaaattcctgccgtcgatgccgagccgttcgccatcttcgtcaat 456
Db 327 TGCCGCGCGGTGGTCGAGGACACGGTGAAACAGCTTGGCGGACTGGACATCGTCGTCAAC 268

Qy 457 gcgatggacaccaatccgctggctgccgaccctacggtcattatcaaagaagccgccgag 516
Db 267 AACGTGGCATTCCAGTCGCCGCAGGACGACCTCACCGAGATCCGCCCAGAAGACGCCGAG 208

Qy 517 g 517
Db 207 G 207

```
RESULT  42
AAF88318/c
ID   AAF88318 standard; DNA; 828 BP.
XX
AC   AAF88318;
XX
DT   28-AUG-2001  (first entry)
XX
DE   S. spinosa DNA fragment encoding ORF1, SEQ ID 7.
XX
KW   Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
KW   spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW   macrolide; insecticidal; O-methyltransferase; ds.
XX
OS   Saccharopolyspora spinosa.
XX
PN   DE19957268-A1.
XX
PD   08-MAR-2001.
XX
PF   29-NOV-1999;   99DE-1057268.
XX
PR   27-AUG-1999;   99DE-1040596.
XX
PA   (FARB ) BAYER AG.
XX
PI   Eberz G,  Moehrle V,  Froede R,  Velten R,  Salas JA;
XX
DR   WPI; 2001-267102/28.
DR   P-PSDB; AAB70948.
XX
PT   New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
PT   recombinant production of insecticidal spinosyns and their derivatives
PT   -
XX
PS   Claim 7a; Page 102-104; 354pp; German.
XX
CC   This invention describes a novel method nucleic acid (I) and its encoded
CC   polypeptide (II) containing at least one region that encodes an enzymatic
CC   activity involved in biosynthesis of spinosyns. (I) are used (i) to
CC   identify, inactivate or modulate genes involved in the biosynthesis of
CC   (II); (ii) to generate a library of polyketide synthases; (iii) for
CC   adding forosamine or trimethylrhamnose to a spinosyn or polyketide
CC   aglycone; and (iv) for recombinant production of the corresponding
CC   enzymes, which are used for production of (II), their precursors or
CC   derivatives, including production of transgenic plants that express (II)
CC   and thus have increased resistance to insects. (I) are also useful as
CC   markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
CC   macrolides with insecticidal, but not antibacterial, activity, and can
CC   also be used to raise specific antibodies, useful for identifying
CC   expression clones in a gene bank. Cells transformed with (I) may produce
CC   (II) at significantly increased levels or produce new derivatives of
CC   (II). This sequence, ORF 1, encodes an S. spinosa O-methyltransferase
CC   protein.
XX
SQ   Sequence 828 BP; 140 A; 262 C; 285 G; 141 T; 0 other;

  Query Match             2.6%;  Score 34.6;  DB 22;  Length 828;
  Best Local Similarity   56.6%;  Pred. No. 4;
  Matches   64;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
           | ||| ||||   ||  ||||||| |  |   |||  || |  || |  |   | ||||
Db     197 CCGCCATCGAGCACGGTCCGTTCGGCGACAAGGTCGGTGAGCCGGTCGGCGGCCTGCTGC 138

Qy     479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggc 531
           | |    |||  | ||| || ||  || |||||  | ||||  |  | | | |
Db     137 CAGGAAGCCCGCCCGTCGTTCTCCCAGTAGCCGTGGTGGATGGCGCAGGGGCC 85

RESULT  43
AAF88315
ID   AAF88315 standard; DNA; 45624 BP.
XX
AC   AAF88315;
XX
DT   28-AUG-2001  (first entry)
XX
DE   S. spinosa DNA fragment SEQ ID 4.
XX
KW   Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
KW   spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW   macrolide; insecticidal; ds.
XX
OS   Saccharopolyspora spinosa.
XX
PN   DE19957268-A1.
XX
PD   08-MAR-2001.
XX
PF   29-NOV-1999;   99DE-1057268.
XX
PR   27-AUG-1999;   99DE-1040596.
XX
PA   (FARB ) BAYER AG.
XX
PI   Eberz G,  Moehrle V,  Froede R,  Velten R,  Salas JA;
XX
DR   WPI; 2001-267102/28.
XX
PT   New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
PT   recombinant production of insecticidal spinosyns and their derivatives
PT   -
XX
PS   Claim 7; Page 58-74; 354pp; German.
XX
CC   This invention describes a novel method nucleic acid (I) and its encoded
CC   polypeptide (II) containing at least one region that encodes an enzymatic
CC   activity involved in biosynthesis of spinosyns. (I) are used (i) to
CC   identify, inactivate or modulate genes involved in the biosynthesis of
CC   (II); (ii) to generate a library of polyketide synthases; (iii) for
CC   adding forosamine or trimethylrhamnose to a spinosyn or polyketide
CC   aglycone; and (iv) for recombinant production of the corresponding
CC   enzymes, which are used for production of (II), their precursors or
CC   derivatives, including production of transgenic plants that express (II)
CC   and thus have increased resistance to insects. (I) are also useful as
CC   markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
CC   macrolides with insecticidal, but not antibacterial, activity, and can
CC   also be used to raise specific antibodies, useful for identifying
CC   expression clones in a gene bank. Cells transformed with (I) may produce
CC   (II) at significantly increased levels or produce new derivatives of
CC   (II). This sequence represents a genomic DNA fragment of the S. spinosa
CC   genome which contains the coding regions for proteins involved in
CC   forosamine and trimethylrhamnose biosynthesis.
XX
SQ   Sequence 45624 BP; 7933 A; 14913 C; 14940 G; 7838 T; 0 other;

  Query Match             2.6%;  Score 34.6;  DB 22;  Length 45624;
  Best Local Similarity   56.6%;  Pred. No. 28;
  Matches   64;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
           | ||| ||||   ||  ||||||| |  |   |||  || |  || |  |   | ||||
Db     632 ccgccatcgagcacggtccgttcggcgacaaggtcggtgagccggtcggcggcctgctgc 691

Qy     479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggc 531
           | |    |||  | ||| || ||  || |||||  | ||||  |  | | | |
Db     692 caggaagcccgcccgtcgttctcccagtagccgtggtggatggcgcaggggcc 744

RESULT  44
AAF88312/c
ID   AAF88312 standard; DNA; 50000 BP.
```

```
XX
AC   AAF88312;
XX
DT   28-AUG-2001  (first entry)
XX
DE   S. spinosa DNA fragment SEQ ID 1.
XX
KW   Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
KW   spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW   macrolide; insecticidal; ds.
XX
OS   Saccharopolyspora spinosa.
XX
PN   DE19957268-A1.
XX
PD   08-MAR-2001.
XX
PF   29-NOV-1999;   99DE-1057268.
XX
PR   27-AUG-1999;   99DE-1040596.
XX
PA   (FARB ) BAYER AG.
XX
PI   Eberz G,  Moehrle V,  Froede R,  Velten R,  Salas JA;
XX
DR   WPI; 2001-267102/28.
XX
PT   New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
PT   recombinant production of insecticidal spinosyns and their derivatives
PT   -
XX
PS   Claim 7; Page 14-31; 354pp; German.
XX
CC   This invention describes a novel method nucleic acid (I) and its encoded
CC   polypeptide (II) containing at least one region that encodes an enzymatic
CC   activity involved in biosynthesis of spinosyns. (I) are used (i) to
CC   identify, inactivate or modulate genes involved in the biosynthesis of
CC   (II); (ii) to generate a library of polyketide synthases; (iii) for
CC   adding forosamine or trimethylrhamnose to a spinosyn or polyketide
CC   aglycone; and (iv) for recombinant production of the corresponding
CC   enzymes, which are used for production of (II), their precursors or
CC   derivatives, including production of transgenic plants that express (II)
CC   and thus have increased resistance to insects. (I) are also useful as
CC   markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
CC   macrolides with insecticidal, but not antibacterial, activity, and can
CC   also be used to raise specific antibodies, useful for identifying
CC   expression clones in a gene bank. Cells transformed with (I) may produce
CC   (II) at significantly increased levels or produce new derivatives of
CC   (II). This sequence represents a genomic DNA fragment of the S. spinosa
CC   genome which contains the coding regions for proteins involved in
CC   forosamine, trimethylrhamnose and polyketide synthase biosynthesis.
XX
SQ   Sequence 50000 BP; 8484 A; 16290 C; 16535 G; 8691 T; 0 other;


  Query Match             2.6%;  Score 34.6;  DB 22;  Length 50000;
  Best Local Similarity   56.6%;  Pred. No. 29;
  Matches   64;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
           | ||| ||||   ||  |||||||  |   |    |||  || |  || |   |   |||||
Db   44993 CCGCCATCGAGCACGGTCCGTTCGGCGACAAGGTCGGTGAGCCGGTCGGCGGCCTGCTGC 44934

Qy     479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggc 531
           | |    |||  | ||| || ||  || |||||  | ||||  |  | | | |
Db   44933 CAGGAAGCCCGCCCGTCGTTCTCCCAGTAGCCGTGGTGGATGGCGCAGGGGCC 44881


RESULT  45
AAZ21501/c
ID   AAZ21501 standard; DNA; 80161 BP.
XX
AC   AAZ21501;
XX
DT   01-DEC-1999  (first entry)
XX
DE   DNA fragment of Saccharopolyspora spinosa containing biosynthetic genes.
XX
KW   Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
KW   microlides; arachnid; nematode; insect; polyketide; polyketide synthase;
KW   PKS; extender module; initiator module; acyl transferase domain; AT;
KW   acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
KW   dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
KW   insecticide; ss.
XX
OS   Saccharopolyspora spinosa.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (1135..1971)
FT                   /*tag=  a
FT                   /product= ORFL16
FT                   /note= "Protein involved in transcription control"
FT   CDS             2024..2791
FT                   /*tag=  b
FT                   /product= ORFL15
FT                   /note= "Keto acyl reductase"
FT   CDS             complement (3416..4165)
FT                   /*tag=  c
FT                   /product= spnS
FT                   /note= "Spinosyn biosynthesis protein S"
FT   CDS             complement (4168..5325)
FT                   /*tag=  d
FT                   /product= spnR
FT                   /note= "Spinosyn biosynthesis protein R"
FT   CDS             complement (5363..6751)
FT                   /*tag=  e
FT                   /product= spnQ
FT                   /note= "Spinosyn biosynthesis protein Q"
FT   CDS             7083..8450
FT                   /*tag=  f
FT                   /product= spnP
FT                   /note= "Spinosyn biosynthesis protein P"
FT   CDS             8967..10427
FT                   /*tag=  g
FT                   /product= SpnO
FT                   /note= "Spinosyn biosynthesis protein O"
FT   CDS             10436..11434
FT                   /*tag=  h
FT                   /product= spnN
FT                   /note= "Spinosyn biosynthesis protein N"
FT   CDS             complement (11530..12492)
FT                   /*tag=  i
FT                   /product= spnM
FT                   /note= "Spinosyn biosynthesis protein M"
FT   CDS             complement (12696..13547)
FT                   /*tag=  j
FT                   /product= spnL
FT                   /note= "Spinosyn biosynthesis protein L"
FT   CDS             complement (13592..14785)
FT                   /*tag=  k
FT                   /product= spnK
FT                   /note= "Spinosyn biosynthesis protein K"
FT   CDS             complement (14799..16418)
FT                   /*tag=  l
FT                   /product= spnJ
FT                   /note= "Spinosyn biosynthesis protein J"
FT   CDS             16556..17743
FT                   /*tag=  m
FT                   /product= spnI
FT                   /note= "Spinosyn biosynthesis protein I"
FT   CDS             complement (17749..18501)
FT                   /*tag=  n
FT                   /product= spnH
FT                   /note= "Spinosyn biosynthesis protein H"
FT   CDS             complement (18541..19713)
FT                   /*tag=  o
```

```
FT                   /product= spnG
FT                   /note= "Spinosyn biosynthesis protein G"
FT   CDS             20168..20995
FT                   /*tag= p
FT                   /product= spnF
FT                   /note= "Spinosyn biosynthesis protein F"
FT   CDS             21111..28898
FT                   /*tag= q
FT                   /product= spnA
FT                   /note= "Spinosyn biosynthesis protein A"
FT                   /transl_except= (Pos:26940..26942, aa:Pro)
FT   misc_feature    21126..22379
FT                   /*tag= r
FT                   /note= "Beta-ketosynthase domain (KSi): part of the
FT                   initiator module"
FT   misc_feature    22692..23669
FT                   /*tag= s
FT                   /note= "Acyl transferase domain (AT1): part of the
FT                   initiator module"
FT   misc_feature    23793..24041
FT                   /*tag= t
FT                   /note= "Acyl carrier protein domain (ACPi): part of the
FT                   initiator module"
FT   misc_feature    24102..25349
FT                   /*tag= u
FT                   /note= "Beta-ketosynthase domain (KS1): part of
FT                   extender module 1"
FT   misc_feature    25683..26684
FT                   /*tag= v
FT                   /note= "Acyl transferase domain (AT1): part of
FT                   extender module 1"
FT   misc_feature    27582..28121
FT                   /*tag= w
FT                   /note= "Beta-ketoreductase domain (KR1): part of
FT                   extender module 1"
FT   misc_feature    28404..28649
FT                   /*tag= x
FT                   /note= "Acyl carrier protein domain (ACP1): part of
FT                   extender module 1"
FT   CDS             38916..35374
FT                   /*tag= y
FT                   /product= spnC
FT                   /note= "Spinosyn biosynthesis protein C"
FT   misc_feature    29024..30295
FT                   /*tag= z
FT                   /note= "Beta-ketosynthase domain (KS2): part of
FT                   extender module 2"
FT   misc_feature    30629..31621
FT                   /*tag= aa
FT                   /note= "Acyl transferase domain (AT2): part of
FT                   extender module 2"
FT   misc_feature    31697..32254
FT                   /*tag= ab
FT                   /note= "Dehydratase domain (DH2): part of extender
FT                   module 2"
FT   misc_feature    33035..34072
FT                   /*tag= ac
FT                   /note= "Enoyl reductase domain (ER2): part of
FT                   extender module 2"
FT   misc_feature    34082..34621
FT                   /*tag= ad
FT                   /note= "Beta-ketoreductase domain (KR2): part of
FT                   extender module 2"
FT   misc_feature    34886..30295
FT                   /*tag= ae
FT                   /note= "Acyl carrier protein domain (ACP2): part of
FT                   extender module 2"
FT   CDS             35419..44931
FT                   /*tag= af
FT                   /product= spnC
FT                   /note= "Spinosyn biosynthesis protein C"
FT   misc_feature    35518..36786
FT                   /*tag= ag
FT                   /note= "Beta-ketosynthase domain (KS3): part of
FT                   extender module 3"
FT   misc_feature    37108..38097
FT                   /*tag= ah
FT                   /note= "Acyl transferase domain (AT3): part of
FT                   extender module 3"
FT   misc_feature    38992..39528
FT                   /*tag= ai
FT                   /note= "Beta-ketoreductase domain (KR3): part of
FT                   extender module 3"
FT   misc_feature    39790..40035
FT                   /*tag= aj
FT                   /note= "Acyl carrier protein domain (ACP3): part of
FT                   extender module 3"
FT   misc_feature    40102..41373
FT                   /*tag= ak
FT                   /note= "Beta-ketosynthase domain (KS4): part of
FT                   extender module 4"
FT   misc_feature    41713..42705
FT                   /*tag= al
FT                   /note= "Acyl transferase domain (AT4): part of
FT                   extender module 4"
FT   misc_feature    43615..44157
FT                   /*tag= am
FT                   /note= "Beta-ketoreductase domain (KR4): part of
FT                   extender module 4"
FT   misc_feature    44431..44676
FT                   /*tag= an
FT                   /note= "Acyl carrier protein domain (ACP4): part of
FT                   extender module 4"
FT   CDS             44966..59752
FT                   /*tag= ao
FT                   /product= spnD
FT                   /note= "Spinosyn biosynthesis protein D"
FT   misc_feature    45077..46348
FT                   /*tag= ap
FT                   /note= "Beta-ketosynthase domain (KS5): part of
FT                   extender module 5"
FT   misc_feature    46691..47674
FT                   /*tag= aq
FT                   /note= "Acyl transferase domain (AT5): part of
FT                   extender module 5"
FT   misc_feature    47753..48310
FT                   /*tag= ar
FT                   /note= "Dehydratase domain (DH5): part of extender
FT                   module 5"
FT   misc_feature    49226..49771
FT                   /*tag= as
FT                   /note= "Beta-ketoreductase domain (KR5): part of
FT                   extender module 5"
FT   misc_feature    50009..50254

  Query Match             2.6%;  Score 34.6;  DB 20;  Length 80161;
  Best Local Similarity   56.6%;  Pred. No. 37;
  Matches   64;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     419 ctgccgtcgatgccgagccgttcgccatcttcgtcaatgcgatggacaccaatccgctgg 478
           | ||| ||||   ||  ||||||| |  |   |||  || |  || |  |   | ||||
Db   20364 CCGCCATCGAGCACGGTCCGTTCGGCGACAAGGTCGGTGAGCCGGTCGGCGGCCTGCTGC 20305

Qy     479 ctgccgaccctacggtcattatcaaagaagccgccgaggatttcaaacgcggc 531
           | |    |||  | ||| || ||  || |||||  | ||||  |  | | | |
Db   20304 CAGGAAGCCCGCCCGTCGTTCTCCCAGTAGCCGTGGTGGATGGCGCAGGGGCC 20252


Search completed: June 30, 2002, 07:12:43
Job time: 43796 sec
```